Third-Party Intervention in Conflicts and the Indirect Samaritan's Dilemma*

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

First Version: June 11, 2009
This Version: July 31, 2009

Abstract

I study a two-period model of conflict 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 between the third party and her ally's enemy where preferences do not satisfy the single-crossing condition. There exist perfect Bayesian equilibria in which the third party's intervention worsens the conflict by energizing her ally's enemy wherein he (i.e., the enemy) pretends to be stronger than he actually is in order to discourage the third-party from assisting her ally. This creates a dilemma for the third party which may be referred to as the *indirect* Samaritan's dilemma. I find that the expectation of a third-party's military assistance to an ally coupled with the third-party's limited information about the strength of her ally's enemy can be strategically exploited by the enemy through *pronouncements* that would not have been credible if the third party was fully informed about her ally's enemy. Remarkably, the third-party's ally, who is fully informed about the enemy, is unable to counteract this behavior by using credible signals to reveal his information to the third party. In some cases, the third party and her ally are strictly better off if the third-party's decision to withdraw from or stay in the conflict is based on her prior beliefs and not on the current conditions of the conflict even if observing the current conditions improves the third-party's information. Unlike the standard Samaritan's dilemma, a commitment by the third party to a given level of assistance may be welfare-improving.

Keywords: Bayesian equilibrium, Grossman-Perry refinement, conflict, intuitive criterion, Samaritan's dilemma.
JEL Classification: D72, D74.
*I thank Bram Cadsby, Rene Kirkegaard, Osman Ouattara, and Jun Zhang for helpful comments. An earlier version of this paper was circulated under the title "Third party intervention in conflicts: feigning weakness and the indirect Samaritan's dilemma."

1. Introduction

The literature on conflicts is large and growing.[1] Within this literature, research on third-party intervention in conflicts is fairly new and the volume is small. In the post-world war II era and the end of the cold war, there have been numerous third-party interventions in conflicts. Between 1944 and 1999, Regan (2002) identified 150 intrastate conflicts of which 101 had third-party interventions. Third-party interventions in conflicts have been in places such as Bosnia, Somalia, Haiti, the former Soviet republics, and Cambodia. These interventions have involved countries like Britain, China, France, and USA and international organizations like the UN.

Regan and Stam (2000) undertook an empirical analysis of third party intervention in conflicts and found that interventions in the earlier stages of a conflict are more effective. 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 Watkins and Winters (1997) argued that biased interventions may be desirable.

The issue of third-party intervention in conflicts has recently received some attention from economists using formal game-theoretic models. Amegashie and Kutsoati (2007) endogenized a third-party's choice of her ally while Carment and Rowlands (1998), Siqueria (2003), and Chang, Potter, and Sanders (2007) took the third-party's ally

[1] See Blattman and Miguel (2009) and Collier and Hoeffler (2007) for surveys of the literature.

as given and examined the effect of third-party's intervention on conflicts. Unlike the present paper, none of these papers incorporate incomplete information amongst the parties.

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. There is a signaling game between the third party party and the enemy in period 1 where the enemy signals his type through the choice of his armed investments. A striking result is that there are perfect Bayesian equilibria in which the third-party's intervention worsens the conflict by inducing the enemy of the third-party's ally to pretend to be stronger than he actually is in period 1 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.

Interestingly, I find that the expectation of a third-party's military assistance to an ally coupled with the third-party's limited information about the strength of her ally's enemy can be strategically exploited by the enemy through *pronouncements* that are credible. Such pronouncements would not have been credible if the third party had full information about her ally's enemy. To be precise, when faction A (i.e., the third-party's ally) is the first mover in the conflict, faction B (i.e., his enemy), who is the second mover, reacts to faction A's investment in arms if the third party has full information about faction B's strength. But when the third party does not have full information about faction B, then faction B can credibly announce that he will choose an armed investment in the current period greater than his full-information investment, regardless of what faction A does in the current period. This announcement has credibility because although it is costly to faction B in the current period, it is beneficial to him in the future because the third-party's belief that he is stronger than he actually is causes the third party to reduce her military assistance to her ally.[2] If faction B had chosen his full-information investment, the third-party would have known his type. Remarkably faction A, who is fully informed about faction B's type, is unable to counteract this behavior by sending credible signals to his ally (the third party) about faction B's type.

It turns out that there is a set of pooling equilibria with such credible bravado that survives the Cho-Kreps "intuitive criterion" but only the Pareto-dominant pooling equilibrium survives the Grossman-Perry refinement. In pooling equilibria, faction B benefits from strategic ambiguity about his type while there is no ambiguity about his actions (i.e., his armed investment). In an interesting paper, Baliga and Sjöström (2008) show that a faction in a conflict can benefit from strategic ambiguity about whether he is armed (i.e., his action).[3]

The result that the enemy of the third-party's ally shows some bravado accords with the intuition that a third-party's intervention could energize the enemy of her ally. However, this is not because the third-party's intervention angers the ally's enemy as intuition might suggest. The enemy's reaction is purely strategic; it has nothing to do with emotions.

[2] This is because the return to the third-party's assistance is smaller, the stronger is faction B. For example, when the USA suffers too many casualities in Iraq, she ponders more over whether her mission is worthwhile. Of course, the USA is directly involved in the conflict in Iraq but there is still the perception that she is there to help the Iraqi government to fight the "insurgents". As I show in appendix C, my results still hold if the third party is directly involved in the conflict.

[3] If he is not armed, he may benefit from not disclosing this information because creating doubt about whether he is armed deters his enemy and if he is armed, he still may not disclose this information because otherwise his enemy might attack him.

My analysis is related to Buchanan's (1975) well-known "Samaritan's dilemma" where a Samaritan's charitable transfers could lead to a perverse behavior by his beneficiary and a commitment not to make such transfers is welfare-improving.

Most models of the Samaritan's dilemma are complete-information models.[4] By incorporating signaling into a model of the Samaritan's dilemma, I follow an approach taken by Lagerlöf (2004), although my application and results are different from his. In Lagerlöf (2004), signaling in a Samaritan's dilemma enhances efficiency while in my model, this need not be the case. While the signaling game in Lagerlof (2004) is a standard game in the sense that it satisfies the single-crossing property, the signaling game in this paper does not satisfy this property.[5] Moreover, there are no pooling equilibria in Lagerlöf (2004) that survive the "intuitive criterion" while there are pooling equilibria in my model that survive this refinement and even survive the stronger refinement in Grossman and Perry (1986).

My model differs from the standard Samaritan's dilemma in one respect: in the version of my model with complete information, there is no Samaritan's dilemma because the third-party's assistance does not lead to any moral hazard behavior. It is incomplete information that results in a Samaritan's dilemma. Furthermore, unlike the standard Samaritan's dilemma, a commitment to refuse assistance does not necessarily improve welfare in my model. An interesting implication of the analysis is that what may matter is not a commitment to refuse assistance but instead a commitment by the third party to be strategically ignorant of the current conditions of the conflict and instead base

[4]Previous analyses of the Samaritan's dilemma include Bernheim and Stark (1988), Lindbeck and Weibull (1988), Bruce and Waldman (1990), Coate (1995), and Andolfatto (2002). Within the context of *non-military* foreign aid, Pedersen (2001) and Blouin and Pallage (2009) examine the Samaritan's dilemma.
[5] For examples of signaling games that do not satisfy the single-crossing condition, see Bernheim and Severinov (2003) and Andreoni and Bernheim (2009).

her future military assistance (i.e., to withdraw from or stay in the conflict) on her prior beliefs. For example, if the equilibrium is a pooling equilibrium, then the third party and her ally are better off if the third party chooses to be strategically ignorant of the conflict in the current period. This is because the third party does not improve her information in a pooling equilibrium and the enemy of the third-party's ally will choose his lower full-information investment in arms if the third party could commit to being strategically ignorant of the conflict in the current period[6] and instead base her assistance on her prior beliefs. Therefore, unlike the standard Samaritan's dilemma, a commitment by the third party to a given level of assistance may be welfare improving. An obvious implication is that if the third party is not fully informed about her ally's enemy, then the fact that the current situation of the conflict looks very bad should not necessarily be used as the basis for withdrawing from the conflict. This is consistent with the view of George W. Bush and Dick Cheney on keeping American troops in Iraq.

However, the welfare implication of such strategic ignorance and commitment is ambiguous if the equilibrium is a separating one because although the stronger type of the enemy may pretend to be stronger than he is in the current period, the third party can make a more informed decision in the future about her military assistance to her ally. If the benefit of the latter effect dominates the cost of the former effect, then strategic ignorance and therefore a commitment to a given level of military support regardless of the conditions on the ground is not welfare improving. In this case, a commitment to withdrawing military support may be welfare improving. Yet, it is also possible that the

[6] For an argument that strategic ignorance can be beneficial in an entirely different context, see Carrillo and Mariotti (2000).

third party and her ally may be better off if she can commit to ignoring any improved information that she may have.

It should be clear from the preceding discussions that even if the third party cannot commit to being strategically ignorant of the conflict in cases where such strategic ignorance would be welfare improving, what ultimately matters is a commitment to discard whatever improved information is acquired and instead commit to a given level of assistance based on the third-party's prior beliefs. Alternatively, the intervention could be designed to be unexpected, although in most cases, this may not be easy.

The paper's main result may be referred to as the *indirect* Samaritan's dilemma. This is because the third-party's assistance does not cause her beneficiary to engage in undesirable behavior. However, her assistance causes those that are in strategic relationships with her beneficiary to engage in undesirable behavior. This indirect Samaritan's dilemma effect is not present in previous analysis of the Samaritan's dilemma because the beneficiary of the Samaritan's charity is not engaged in any strategic relationship with other parties. The indirect Samaritan's dilemma may have implications for other applications apart from the model of conflict studied here. I discuss an application in my concluding remarks.

My result does not contradict the aforementioned empirical finding of Regan and Stam (2000) because I am not arguing that actual military support *per se* increases the incidence of conflict. My argument is that the *expectation* of military support could increase the incidence of conflict. For example, in my two-period model, the incidence of conflict in period 1 would be reduced if the third party could offer military support in period 1. It is the *expectation* of military support in period 2 coupled with the third-party limited information about her ally's enemy that increases the incidence of conflict in period 1. This has implications for empirical work. It suggests that conflicts may worsen prior to a publicly known and biased intervention and will 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 reduction in the intensity of the conflict is not necessarily due to the 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).

The preceding argument somewhat finds support in the empirical work of Elbadawi and Sambanis (2000). 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. Elbadawi and Sambanis (2000) found that external intervention increased the duration of conflict.[7]

My paper also contributes to the new and small literature on signaling in contests. Hörner and Sahuguet (2007), Munster (2009), Zhang and Wang (2009), and Slantchev (2009) examine dynamic contests with signaling. However, in these papers, the signaling game is between the players in the contest and there is no third party. In my model, signaling game is between a player in the contest and an outside player (i.e., the third party).

[7] Of course, caution must be exercised here because my model is about the intensity of conflict and not about its duration. But more importantly, my argument for using expected intervention is different from Elbadawi and Sambanis (2000) because 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.

In Munster (2009) and Slantchev (2009), where the combatants move simultaneously in the conflict, they show that a strong player may feign weakness by pretending to be weaker than he really is in order to surprise his opponent in a subsequent period of conflict. This is also possible in my model if say faction A is uninformed about faction B's type and they move simultaneously in the conflict. And in a dynamic auction where a first mover can use a jump bid to signal his type to a second mover, Hörner and Sahuguet (2007) show that both *feigning weakness* and *displaying bravado* are possible equilibria. In their model, displaying bravado is beneficial because it may cause one's opponent to succumb which is similar to the motivation in my paper. And feigning weakness is beneficial because it cause one's opponent to slack in a subsequent round.[8]

However, my paper is different from above papers. First, although Hörner and Sahuguet's (2007) result shows that displaying bravado could exist in conflicts without third-party intervention, my focus on third-party intervention in conflicts is important because it allows me to identify a different and important motivation for why a player in a conflict may display bravado and also to point out a possible perverse effect of a common real-world practice in conflicts (i.e., third-party intervention). Second, these papers do not obtain the result that a *de jure* second mover in a conflict can use his superior information to become a *de facto* first mover. Third, they do not uncover the indirect Samaritan's dilemma effect because there is no third party in their models. Fourth, these papers do not discuss how a third party should react to the current conditions of a conflict in the sense that strategic ignorance may be helpful or the

[8] Zhang and Wang (2009) is different from these papers because they focus on the role of information revelation in contests.

withdrawal decision need not be based on current conditions of the conflict.[9] In short, these papers focus on different issues.

Finally, the paper contributes to the literature on the endogenous timing of moves in contests inspired by Dixit (1987): Baik and Shogren (1992), Leininger (1993), Morgan (2003) and Morgan and Vardy (2007). In particular, Baik and Shogren (1992) and Leininger (1993) show if two players in a contest can choose the order of moves, then , the player with the higher valuation moves second and the player with the lower valuation moves first.[10] This paper shows that, regardless of the players' relative valuations, incomplete information can lead to a *de jure* second mover in a contest acting as a *de facto* first mover.

The paper is organized as follows. The next section presents a model and analysis of signaling in a conflict where I demonstrate what I call the *indirect* Samaritan's dilemma. Several subsections deal with implications of the analysis. Section 3 concludes the paper. There are three appendices that elaborate on certain issues. In particular, appendix A shows that the players' preferences do not satisfy the single-crossing condition and appendix B explains how an important parameter of the third-party's out-of-equilibrium beliefs is determined. Appendix C presents an alternative model of third-party intervention.

[9] In Hörner and Sahuguet (2007), a player may show bravado, but will succumb when his bluff is called. In my model, this is not possible. If the player will succumb in period 2, then he will not show bravado in period 1.

[10] Morgan and Felix (2007) show that this result does not hold if the second mover can only observe the effort chosen by the first mover by incurring a cost.

2. A model of third-party intervention and signaling

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.[11] There are two time periods, 1 and 2. In each period, there could be a fight (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 (country), the third party 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. 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.

In each period, the factions in the conflict 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

[11] One may argue that the third party (e.g., the USA) may be directly involved in the conflict. I consider this case in appendix C. Moreover, it must be noted that third parties do intervene in conflicts by giving financial support for military expenditure without directly getting involved in the conflict.

$P_A^1 = G_A/(G_A + G_B)$ and $P_B^1 = 1 - P_A^1$ of the land (country). A similar function describes the proportions in period 2.

Without loss generality, I assume that the third party can help faction A in period 2 but not in period 1.[12] 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.

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 subsidy transfer 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 subsidy reduces faction A's unit cost of arming and $0 < \theta < 1$. In appendix C, I consider a different formulation where the third party is directly involved in the conflict by choosing the quantity of arms to assist faction A. The results remain unchanged.

The timing of actions is as follows:

Period 1:

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 A chooses his investment in arms

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

[12] 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 is taken before faction B moves. It also means that I could demonstrate my result in one-period model with more stages. However, the two-period model below is more convenient for exposition. In any case, this assumption is the same as an assumption in models of the Samaritan's dilemma where the Samaritan only lives in period 2 but her beneficiary lives in periods 1 and 2 (e.g., see Lagerlöf, 2004).

Period 2:

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.[13]

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

Stage 2: Both factions observe the third-party's choice and faction A chooses his investment in arms.

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

I look for perfect Bayesian Nash equilibria of this game and restrict attention to pure strategies. However, I first begin with the benchmark case of complete information where I look for a subgame perfect Nash equilibrium.

2.1 Benchmark: complete information

The single-period and complete-information version of the above repeated game without third-party intervention has been studied by Grossman and Kim (1995) and Gershenson and Grossman (2000).[14] Chang, Potter, and Sanders (2007) extend this model by introducing third party intervention. In this section, I follow the analysis in these papers.

Consider a single period and assume that the third party can give military assistance to her ally. In particular, consider period 2.

In stage 3, faction B of type k chooses X_B to maximize

$$\Pi_{k|B} = \frac{X_{k|B}}{X_{k|B} + X_{k|A}} W_k - X_{k|B}, \tag{1}$$

[13]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.

[14] See also Leininger (1993) and Morgan (2003).

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:

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

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 $X_{k|A} = W_k$. 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

$$\Pi_{k|A} = \frac{X_{k|A}}{X_{k|B} + X_{k|A}} V - \frac{1}{(1 + M_k)^{\theta}} X_{k|A} , \tag{4}$$

where M_k is the third-party's assistance to faction A if his opponent is of type k, k = H, L.

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

$$\Pi_{k|A} = V \sqrt{\frac{X_{k|A}}{W_k}} - \frac{1}{(1 + M_k)^{\theta}} X_{k|A} \tag{4a}$$

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

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

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

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

Suppose that $M_k = 0$. Then using equation (6), $X^*_{k|B} > 0$ 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. Indeed, a sufficient condition for my analysis is $V < 2W_k$ for, at least, one k. [15]

The equilibrium proportions of the land controlled by faction A conditional on M_k are:

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

The third party will choose M_k to maximize

$$\Pi_{k|C} = P^*_{k|A}S - M_k \tag{9}$$

This gives

$$M^*_k = \left(\frac{\theta S}{2}\frac{V}{W_k}\right)^{1/(1-\theta)} - 1 \tag{10}$$

I assume that $M^*_L > 0$ but $M^*_H \geq 0$. So when faction B is weak, the third party will always intervene but may back off when faction B is strong.

A few remarks are in order. Unlike the standard Samaritan's dilemma, faction A's payoff does not enter the third-party's payoff function. However, there is a component of

[15] If $V \geq W_k$ 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.

faction A's payoff which enters the third-party's payoff function. That is, the third party cares about the proportion of the land controlled by faction A and his payoff is increasing in faction A's investment in arms. Over a certain range of investment in arms, faction A's payoff is also increasing in his investment. Yet, as will be evident below, the expectation of the third-party's assistance may induce faction B to invest more in arms which causes faction A to reduce his investment and reduces the payoffs of faction A and the third party in the current period. Therefore, the structure of the game is not different from the standard Samaritan's dilemma. Yet, in a two-period version of this game with complete information, there is no Samaritan's dilemma because the third party's assistance does not lead to any undesirable behavior. As will be evident in the next section, it is the incompleteness of information that gives a Samaritan's dilemma effect.

2.2 Incomplete information

Note that the game between factions A and B is a sequential game of complete and perfect information while the game between faction B and the third party is a sequential game of perfect but incomplete information. It turns out that the signaling game here is not a standard signaling game in the sense that it does not satisfy the single-crossing condition. However, the proof of this claim, given in appendix A, must wait till part of the equilibrium of the game is characterized.

There is also a signaling game between the third party and faction A because faction A may want to reveal what he does about faction B to the third party. However, I characterize the equilibrium of the game by assuming that faction A does not choose his investment in order to convey what he knows about faction B to the third party.

Thereafter, I show that the assumed behavior for faction A is indeed an equilibrium behavior.

2.2.1 Equilibrium in period 2

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 and so I shall draw on most of the analysis in section 2.1 with appropriate modifications.

From section 2.1, faction B of type k's best response function in period 2 is $X_{k|B} = \sqrt{W_k X_{k|A}} - X_{k|A}$, k = H, L. Therefore, using equation (5), gives faction A's investment in period 2 when his opponent is faction B of type k as[16]

$$\hat{X}_{k|A} = \frac{V^2(1+M)^{2\theta}}{4W_k}, \qquad (11)$$

where M is the third-party's assistance to faction A that will be determined shortly and k = H, L.

Then faction B of type k invests $\hat{X}_{k|B} = \sqrt{W_k \hat{X}_{k|A}} - \hat{X}_{k|A}$. That is, using (6),

$$\hat{X}_{k|B} = \frac{(1+M)^{\theta} V}{2} - \frac{(1+M)^{2\theta} V^2}{4W_k}, \qquad (12)$$

in period 2, conditional on the third-party's assistance, k = H,L.

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

[16] I focus on interior solutions in which case faction A cannot deter faction B with or without military assistance. As explained in the previous section, a sufficient condition for my analysis is that, in the absence of the third party, faction A cannot deter, at least, one type of faction B.

$$\hat{P}_{k|B} = 1 - \frac{(1+M)^{\theta} V}{2W_k}, \tag{13}$$

$k = H, L$.

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

$$\hat{\Pi}^2_{k|B} = \left(1 - \frac{(1+M)^{\theta} V}{2W_k}\right)\left(W_k - \frac{(1+M)^{\theta} V}{2}\right). \tag{14}$$

Clearly, faction B's payoff in (14) is decreasing in the third-party's military assistance to faction A.

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

$$\Omega^2_C = (1-\mu)\frac{V}{2W_L}(1+M)^{\theta} S + S\mu \frac{V}{2W_H}(1+M)^{\theta} - M \tag{15}$$

The third-party's optimal military assistance is:

$$\hat{M}(\mu) = \left[\frac{\theta SV}{2}\left(\frac{\mu}{W_H} + \frac{1-\mu}{W_L}\right)\right]^{1/(1-\theta)} - 1 \tag{16}$$

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

I focus on interior solutions by assuming that, in equilibrium the expressions in (12), (13), and (14) are positive. This means that faction A cannot deter either type of faction B even with the third-party's assistance. None of my results hinge on this assumption.

Given (16), 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 (see equation (15)). Therefore, this result and equation (14) 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 pretend to be stronger than he actually is. 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.

2.2.2 Pooling equilibrium in period 1

Consider period 1. Recall that in this period I assume, without any loss of generality, that the third party cannot, give any military assistance. Consider a pooling equilibrium where both types of faction B choose $G_{L|B} = G_{H|B} = \hat{G}_B > G^*_{H|B}$ with $\mu = q$. In this case, faction A's first-mover advantage is useless because it is as though, in period 1, faction B has publicly made it known to faction A that he is committed to an investment of $\hat{G}_B$, greater than his full-information investment, no matter what faction A does. Faction B's pronouncement is credible because of the associated benefits to him in period 2 stemming from the reduced military assistance to faction A. If the third party was fully informed about the strength of faction B, this pronouncement would not have been credible because it would not affect the third-party's decision in period 2. This is an important point and I shall prove this can be sustained in equilibrium.

Therefore, as a second mover, faction B's best-response in period 1 to any investment chosen by faction A is $G_{k|B} = \hat{G}_B \; \forall \; G_A$, k = H,L. To reiterate, faction B can credibly say that he will not react to any investment chosen by faction A in period 1 because by committing to $G_{k|B} = \hat{G}_B$, he will benefit from the conflict in period 2 so long as the third party posterior belief that faction B is strong is greater than her prior belief because of faction B's commitment. This leads to a reduction in the third-party's assistance to faction A.

Mindful of faction B's credible pronouncement, faction A best responds to $\hat{G}_B$ by choosing $\hat{G}_A = \sqrt{V\hat{G}_B} - \hat{G}_B$. Therefore, acting as the *de facto* first mover faction B's payoff in period 1 is

$$\hat{\Pi}_k^1(\hat{G}_B) = \frac{\hat{G}_B}{\hat{G}_B + \sqrt{V\hat{G}_B} - \hat{G}_B} W_k - \hat{G}_B = W_k\sqrt{\frac{\hat{G}_B}{V}} - \hat{G}_B \qquad (17)$$

k = H, L.

Suppose $(W_k)^3 > V^2(2W_k - V)$,[17] k = H,L. Then $0.25(W_k)^2/V = \hat{G}^*_{k|B} = \arg\max_{\hat{G}_{k|B}} \hat{\Pi}_k^1 > G^*_{k|B}$, where $G^*_{k|B} = 0.25V(2W_k - V)/W_k > 0$[18] is, of course, the full-information investment of faction B of type k in period 1 (i.e., when the third party has full information in which case, as a second mover, he reacts to faction A's choice of investment). Note that, $\hat{\Pi}_L^1$ is strictly increasing in $\hat{G}_B$ for $\hat{G}_B \in [0, \hat{G}^*_{L|B})$ attaining its maximum at $\hat{G}^*_{L|B}$ and given that $W_H > W_L$, $\hat{\Pi}_H^1$ is strictly increasing in $\hat{G}_B$ for $\hat{G}_B \in [0, \hat{G}^*_{L|B}]$. Also, $\partial\hat{\Pi}_H^1/\partial\hat{G}_B > \partial\hat{\Pi}_L^1/\partial\hat{G}_B$ at any given $\hat{G}_B$.

[17] This condition is not crucial. If it is not satisfied we can still construct pooling equilibria with *bravado* that survive the intuitive criterion but not the Grossman-Perry refinement.
[18] One obtains this equation by setting M = 0 in equation (12).

Finally, it is useful to note that $\hat{\Pi}_k^1$ is decreasing in $\hat{G}_B$ for $\hat{G}_B \in [\hat{G}^*_{H|B}, V]$ and $\hat{G}^*_{H|B} >$

$\hat{G}^*_{L|B}$, k = H,L.

Before I continue, an important remark is in order. Nothing in the construction of the equilibria depend on faction B's *de facto* first-mover role. The pooling and separating equilibria below still hold by assuming that faction B is the *de jure* first mover.[19] The interesting aspect about the present set-up is that it allows us to argue that faction B can exploit his informational advantage vis-à-vis the third party to act in a way which makes him the *de facto* first mover even though faction A is the *de jure* first mover.

Conditional on $\hat{G}_A$, consider the following out-of-equilibrium beliefs for the third party:

$$\mu(G_B) \equiv \mu(G_B, \hat{G}_A) = \begin{cases} 1 & \text{if} \quad G_B \in (\hat{G}_B + \lambda, \infty) \\ q & \text{if} \quad G_B \in [\hat{G}_B, \hat{G}_B + \lambda] \\ 0 & \text{if} \quad G_B \in [0, \hat{G}_B) \end{cases} \tag{18a}$$

where $\lambda > 0$. In appendix B, I explain how λ is determined.

Notice that, following a standard notion of a Nash equilibrium, I do not allow joint deviations from equilibrium by factions A and B. Therefore, the third-party's out-of-equilibrium beliefs about faction B's actions are conditional on faction A playing his (i.e., faction A) equilibrium strategy. Similarly, as indicated below, the third-party's out-of-equilibrium beliefs about faction A's actions are conditional on faction B playing his (i.e., faction B) equilibrium strategy.

---

[19]The only difference is if faction B was the *de jure* first mover in period 1, then his full-information investment will be higher and so in the unique pooling equilibrium one type does not show bravado. In particular, the strong type does not but the weak type does.

For the sake of argument, suppose $\lambda = 0$. Then, if either type of faction B deviates in period 1 by increasing his investment by a very small amount his payoff in period 1 only falls marginally or may increase while his payoff in period 2 increases discontinuously. Under these conditions, both types of faction B will deviate. However, given that it is a strictly dominant strategy for both types to increase their investment by a very small amount, the third party, using Bayes' rule, should believe that $\mu = q$ if she observes very small increases from the equilibrium investment. Hence, the beliefs in (18a) are not reasonable if $\lambda = 0$. This explains why $\lambda > 0$.

Note that given the beliefs in (18a), any $\hat{G}_B \in [0, \hat{G}^*_{H|B})$ cannot be part of a pooling equilibrium. This is because if the strong type of faction B deviates to $\hat{G}^*_{H|B}$, he will increase in his payoff in period 1 because he will be choosing his optimal investment in arms and his payoff in period 2 will not fall because the third party's belief that he is strong will stay the same or increase which means that the third party will not increase his military assistance to faction A. Therefore, it makes sense to focus on $\hat{G}_B \in [\hat{G}^*_{H|B}, V]$ as possible candidates for a pooling equilibrium.[20]

Therefore, suppose $\hat{G}_B \in [\hat{G}^*_{H|B}, V]$. Consider any type of faction B. If he deviates by increasing his investment he will reduce his payoff in period 1 but may increase his payoff in period 2 because the third party will reduce his assistance to faction A if the increase is greater than λ. If the former effect is costlier than the benefit of the latter effect, then he will not deviate. On the other hand, if he deviates by reducing his

[20]If faction B invests V in arms, then faction A will invest zero, so there is no point for faction B to invest more than V.

investment, he may increase his payoff in period 1 if he is a weak type and will reduce his payoff in period 1 if he is a strong type. However, he will reduce his payoff in period 2 because the third party will increase his assistance to faction A. He will not deviate if the latter effect dominates the former effect. Under these conditions, $\hat{G}_B \in [\hat{G}^*_{H|B}, V]$ is a perfect Bayesian pooling equilibrium.

Let $\Theta_k(G_B,\mu) = \hat{\Pi}^1_k(G_B) + \hat{\Pi}^2_k(\hat{M}(\mu))$. Then, given $\hat{G}^1_B \in [G^*_{H|B}, V]$ and the beliefs in (18a), there exists a perfect Bayesian pooling equilibrium if

$$\Theta_k(\hat{G}_B, q) \geq \Theta_k(G_B, \mu(G_B)) \tag{19}$$

for all G_B and k = H,L. It is easy to see that given the beliefs in (18a) and choice of λ in appendix B, we can choose the other parameters of the model to ensure that the inequality in (19) holds.

I use the Cho-Kreps "intuitive criterion" to place restrictions on out-of-equilibrium beliefs. In my model, 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., μ = 1). Accordingly, the pooling equilibria above satisfy the Cho-Kreps "intuitive criterion" if there is no G_B such that

$$\Theta_k(G_B, 1) > \Theta_k(\hat{G}_B, q) \text{ and } \Theta_j(G_B, 1) < \Theta_j(\hat{G}_B, q), \tag{20}$$

k, j = H,L and k $\neq$ j.

Given (19) and the beliefs in (18a), neither type of faction B has an out-of-equilibrium increase in investment greater than or equal to λ that gives a higher payoff than his equilibrium payoff even if the third party believed that the deviation was by a strong type. However, as argued above, there are increases from the equilibrium investment smaller than λ which are strictly dominant strategies for each type if the third party believed that such deviations were made by the strong type. But in such cases, the "intuitive criterion" does not tell us what to do and so has no bite.[21]

Recall that $\partial\hat{\Pi}_H^1 / \partial\hat{G}_B > \partial\hat{\Pi}_L^1 / \partial\hat{G}_B$ at any given $\hat{G}_B$ and $\hat{G}_{H|B}^* > \hat{G}_{L|B}^*$. Therefore, a reduction in investment in period 1 is more costly for the strong type than for the weak type and, in some cases, is even beneficial to the weak type. Also from (11), $\partial\hat{X}_{L|A} / \partial M > \partial\hat{X}_{H|A} / \partial M$. Hence, for a given reduction in military assistance, faction A reduces his investment in period 2 by a bigger amount if faction B is weak than if faction B is strong. Hence the weak type of faction B benefits more from a given reduction in assistance to faction A. Suppose the strong type finds it profitable to deviate to a smaller investment if the third party believed that he was strong. Then the preceding arguments

[21]It is in these cases that stronger refinements like the D1 condition in Cho and Kreps (1987), *divinity* in Banks and Sobel (1987), and the refinement in Grossman and Perry (1986) suggest what to do. The D1 condition would suggest 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) while the *divinity* refinement would suggest that we place *relatively more* weight on this type. In these cases, we can also use the Grossman-Perry refinement. Indeed, as I argue below, the Grossman-Perry refinement gives a unique pooling equilibrium.

imply that the weak type will also find such a deviation profitable.[22] In this case, the "intuitive criterion" does not tell us what to do and so has no bite.

The above arguments imply that any perfect Bayesian pooling equilibrium with $\hat{G}_B \in [\hat{G}^*_{H|B}, V]$ survives the Cho-Kreps "intuitive criterion" and so, based on the intuitive criterion, the beliefs in (18a) are reasonable.

Note that for any $\hat{G}_B \in [\hat{G}^*_{H|B}, V]$, both types of faction B pretend to be stronger than they really are by investing in a higher level of arms relative to the case of complete information or the case of no military support in period 2.

Finally, the analysis has been based on the assumption that conditional on the equilibrium strategies of faction B and the third party, it is optimal for faction A to choose his subgame-perfect investment in each period. In particular, faction A would choose $\hat{G}_A = \sqrt{V\hat{G}_B} - \hat{G}_B \geq 0$ in period 1. However, since faction A knows faction B's type, he may not choose $\hat{G}_A$ but instead choose his investment in period 1 with the goal of revealing faction B's type to the third party. Given $\hat{G}_B$, this strategy is costly to faction A in period 1 but may be beneficial to him in period 2.

Note that faction A has no incentive to signal to the third party that faction B is strong because that means that faction A will get less military assistance. Therefore, his goal is to signal to the third party that faction B is weak. Note also that the third party cannot believe whatever faction A says because he has the incentive to say that faction B

[22]While any decrease (deviation) that the strong type 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 (18a) for investments smaller than the pooling equilibrium investment. This kind of reasoning is in the spirit of the D1 condition of Cho and Kreps (1987).

is weak when in fact he is strong.[23] Then the pooling equilibria above can be supported by arguing that the third party will believe that faction B is weak type if and only if faction A invested a small amount in period 1 which can be suitably chosen to be make it unprofitable for faction A to do so. For example, suppose the third-party's out-of-equilibrium belief is:

$$\mu(G_A) \equiv \mu(\hat{G}_B, G_A) = \begin{cases} 1 & \text{if} \quad G_A \in [\hat{G}_A, \infty) \\ q & \text{if} \quad G_A \in [\hat{G}_A - \gamma, \hat{G}_A] \\ 0 & \text{if} \quad G_A \in [0, \hat{G}_A - \gamma) \end{cases} \tag{18b}$$

where $\gamma > 0$ and $\hat{G}_A - \gamma \geq 0$. Since factions A and B have *diametrically-opposed* objectives in terms of what they want the third party to believe, it is not surprising that the belief in (18b) is a mirror image of the belief in (18a). An analysis similar to the one in appendix B shows how γ should be set to make it unprofitable for faction A to deviate from the equilibrium. Analogously, we can also show that these beliefs satisfy the "intuitive criterion".[24]

The act of bravado by faction B is a dilemma for the third party because her assistance can make things worse in period 1 by energizing faction B in period 1. Yet, in period 2, she cannot commit not to offer assistance to faction A. Therefore, faction B will rationally expect the third party to play her Nash strategy in period 2. This creates what

[23] One may find this counterintuitive on the grounds that faction A will use the difficulty of the war as the basis for asking for help and therefore has the incentive to indeed truthfully report that faction B is strong. But this argument misses a fundamental point. While the third party may appreciate the difficulty of the war, what she is ultimately interested in is if the war can be won (i.e., what is the likelihood of victory). Since for a given assistance, the equilibrium probability of victory or the proportion of the territory controlled by faction A is higher when faction B is weak than when faction B is strong, faction A's incentive to report that faction B is weak is to assure the third party of a high probability of victory or a high return to her assistance in terms of the proportion of territorial control.

[24] These arguments are available on request.

one might call the *indirect* Samaritan's dilemma for the third party for reasons given in section 1.

I summarize the analysis in the following proposition:

Proposition 1 (The indirect Samaritan's dilemma): *There exists a set of pooling equilibria that survive the "intuitive criterion" in which faction B of either type pretends to be stronger than he actually is by choosing an armed investment 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).*

Since the third party's military assistance is the same across all the pooling equilibria, it follows that the pooling equilibrium which gives the highest payoff to either type of faction B is $\hat{G}_B = \hat{G}^*_{H|B}$. It turns out that this is the only pooling equilibrium that satisfies the Grossman-Perry refinement which is stronger than the "intuitive criterion".

To see this, assume that third party believes that a deviator belongs to the set, S, of types. In this case, S = $\{W_H\}$ or $\{W_L\}$ or $\{W_L, W_H\}$. Let the third party update her beliefs using Bayes' rule and choose her military assistance according to $\hat{M}(\mu)$. Grossman and Perry (1986) require the third party to ask the following questions: (a) would all types in S be strictly better off by this deviation given the military assistance chosen by the third party?, and (b) Is it the case that types that do not belong to S will not be better off by this deviation even if they received the same military assistance as those in S? If the answer to (a) and (b) is yes, then the equilibrium fails to satisfy the Grossman-Perry refinement. First, note that if S is a singleton, then the Grossman-Perry refinement is equivalent to the "intuitive criterion". So it remains to apply the refinement to the case of S = $\{W_L, W_H\}$. In this case, the third party maintains her prior beliefs. If so, then both types are strictly better off by deviating to $G_B < \hat{G}_B$ if $\hat{G}_B \in (\hat{G}^*_{H|B}, V)$.

Therefore, these pooling equilibria do not survive the Grossman-Perry refinement. If $\hat{G}_B = G^*_{H|B}$, then the strong type will not deviate but, given $\hat{G}^*_{H|B} > \hat{G}^*_{L|B}$, the weak type will deviate by reducing his investment. Since not all types in $S = \{W_L, W_H\}$ are strictly better off from the deviation, it follows that the pooling equilibrium $\hat{G}_B = G^*_{H|B}$ survives the Grossman-Perry refinement. This leads to the following proposition:

Proposition 2: *Given pooling equilibria in which faction B pretends to be stronger than he actually is, the Grossman-Perry refinement picks the Pareto-dominant pooling equilibrium as the unique equilibrium.*

2.2.3 Separating equilibrium in period 1

In standard signaling games, separating equilibria tend to exist and survive the "intuitive criterion" while pooling equilibria may not exist.[25] Not surprisingly, a separating equilibrium is easier to characterize than a pooling equilibrium in this game. In particular, the choice of out-of-equilibrium beliefs is more straightforward.

In the present game, there is a separating equilibrium in period 1 in which the strong type pretends to be stronger than he is and the weak type chooses his full-information investment, $G^*_{L|B}$. In period 1, the strong type of faction B chooses $\hat{G}^*_{H|B} > G^*_{H|B}$ in period 1 and the weak type chooses $G^*_{L|B}$, where it needs to be recalled

[25] Cho and Sobel (1990) show that standard signaling games (i.e., strategy sets and payoff functions satisfy monotonicity properties) have a separating equilibrium and it is the only equilibrium that satisfies the D1 condition in Cho and Kreps (1987). As would be evident in appendix A, in this signaling game, the informed type's payoff function is not monotonic in his message.

that $\hat{G}^*_{H|B}$ is the investment that maximizes the payoff of the strong type of faction B when he acts likes the *de facto* first mover and $G^*_{H|B}$ is his investment in the full-information case where he is the second mover.[26] The third-party's equilibrium beliefs in period 2 are $\hat{\mu}(G^*_{L|B}) = 0$ and $\hat{\mu}(\hat{G}^*_{H|B}) = 1$. 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 $\hat{M}(0)$.

In equilibrium, faction A chooses $\hat{G}^*_{H|A} = \sqrt{V\hat{G}^*_{H|B}} - \hat{G}^*_{H|B} > 0$ and $G^*_{L|A} = \sqrt{VG^*_{L|B}} - G^*_{L|B} > 0$. Conditional on these equilibrium investments, let the third-party's out-of-equilibrium beliefs be:

$$\hat{\mu}(G_B) = \hat{\mu}(G_B, G^*_{L|A}, \hat{G}^*_{H|A}) = \begin{cases} 1 & \text{if} \quad G_B \in [\hat{G}^*_{H|B}, \infty) \\ \\ 0 & \text{if} \quad G_B \in [0, \hat{G}^*_{H|B}) \end{cases}. \tag{21}$$

Given (21), a separating equilibrium exists if

$$\Theta_H(\hat{G}^*_{H|B}, 1) \geq \Theta_H(G_B, \hat{\mu}(G_B)), \tag{22}$$

and

$$\Theta_L(G^*_{L|B}, 0) \geq \Theta_L(G_B, \hat{\mu}(G_B)) \tag{23}$$

for all G_B. Again, we choose the parameters of model to ensure that (22) and (23) hold.

---

[26] Note that since types are fully revealed in this case, the weak type gains nothing by deviating from his full-information investment. Hence, he best responds to faction A's investment.

I need to argue that this equilibrium is supported by reasonable out-of-equilibrium beliefs. Recall that $\hat{G}^*_{H|B} = \arg\max_{\hat{G}_{H|B}} \hat{\Pi}^1_H > G^*_{H|B}$. Therefore, the strong type of faction B has no incentive to deviate if even if his type is correctly inferred. Therefore, the third party will correctly infer that faction B is weak if she observes any deviation to any $G_B \neq \hat{G}^*_{H|B}$. Hence, the weak type will not deviate to $G_B \neq \hat{G}^*_{H|B}$. But given (23) and the beliefs in (21), the weak type will also not deviate to $G_B = \hat{G}^*_{H|B}$ even if he is mistaken for a strong type. Hence, the weak type will also not deviate.

Now consider faction A. He will never deviate if faction B is indeed weak because in a separating equilibrium, the third party beliefs are correct. So any possible deviation by faction A will occur if and only if faction B is strong. Then in the spirit of the intuitive criterion, we can support the separating equilibrium by using the following simple out-of-equilibrium belief for deviations by faction A: $\hat{\mu}(G_A, G^*_{L|B}, \hat{G}^*_{H|B}) = 1$ if $G_A \neq G^*_{L|A}$. Then faction A will not deviate.

Therefore, it trivially follows that there is a unique separating equilibrium that survives the Grossman-Perry refinement (stronger than the "intuitive criterion") in which the strong type of faction B display some bravado but the weak type does not. Hence, the indirect Samaritan's dilemma also occurs in a separating equilibrium. Note that if $\hat{M}(1) = 0$, then in a separating equilibrium, the third party does not meddle in the conflict if faction B is strong. This gives

Proposition 3 (The indirect Samaritan's dilemma): *There exists a unique separating equilibrium that survives the Grossman-Perry refinement in which the strong type of faction B pretends to be stronger than he actually is by choosing an armed investment 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 full-information investment.*

2.2.4 Welfare

In both the pooling and separating equilibria there is, at least, one type of faction B who is able to nullify faction A's first-mover advantage by committing to an investment level in period 1 precisely because of the third-party's limited information about faction B. Furthermore, the third party can nullify this behavior if she can choose to be strategically rationally ignorant of conflict in period 1. This can be summarized in the following proposition:

Proposition 4: *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 through pronouncements that are credible. However, the ability of the enemy to strategically gain from his superior information no longer exists if the third party can strategically commit to being ignorant of the conflict or if she can commit to a given level of assistance based on her prior beliefs.*

As explained in section 1, these corollaries follow from proposition 4:

Corollary 1: *If the equilibrium in period 1 is a pooling equilibrium, then the third party and her ally are strictly better off if the third party can strategically commit to being ignorant of the conflict in period 1 or if she can commit to a given level of assistance in period 2 based on her prior beliefs.*

Corollary 2: *If the equilibrium in period 1 is a separating equilibrium, then the third party and her ally may be strictly better off if the third party can strategically commit to being ignorant of the conflict in period 1 or if she can commit to a given level of assistance in period 2 based on her prior beliefs. In this case, the third party ignores any improved information that she has.*

The intuition for these corollaries was given in section 1, so I will not rehash here. The reader may refer to the discussion in section 1.

2.2.5 Pooling or separating equilibrium?

We can impose restrictions on the parameters of the model by eliminating one of the equilibria in propositions 2 and 3. Note that the strong type of faction B invests $\hat{G}^*_{H|B}$ in each of these equilibria. Also, in each of these equilibria, the respective out-of-equilibrium beliefs in (18a) and (21) is that the third party believes that faction B is weak if $G_B \in [0, \hat{G}^*_{H|B}]$ conditional on the equilibrium strategy of faction A.

For all $G_B \in [0, \hat{G}^*_{H|B}]$, the inequality in (22) and the beliefs in (21) mean that a necessary condition for a separating equilibrium is

$$\Theta_H(\hat{G}^*_{H|B}, 1) \geq \Theta_H(G_B, 0) \tag{24}$$

and

(18a) and (19) also mean that a necessary condition for a pooling equilibrium is

$$\Theta_L(\hat{G}^*_{H|B}, q) \geq \Theta_L(G_B, 0) \tag{25}$$

for all $G_B \in [0, \hat{G}^*_{H|B}]$.

Then given $\Theta_H(\hat{G}^*_{H|B}, 1) > \Theta_H(\hat{G}^*_{H|B}, q)$, we can violate a necessary condition for the pooling equilibrium in proposition 2 and construct the separating equilibrium in proposition 3 such that

$$\Theta_H(\hat{G}^*_{H|B}, q) < \Theta_H(G_B, 0) \leq \Theta_H(\hat{G}^*_{H|B}, 1) \tag{26}$$

for some $G_B \in [0, \hat{G}^*_{H|B}]$.

Similarly, given $\Theta_L(\hat{G}^*_{H|B}, q) > \Theta_L(\hat{G}^*_{H|B}, 0)$, we can violate a necessary condition for the separating equilibrium in proposition 3 and construct the pooling equilibrium in proposition 2 such that

$$\Theta_L(\hat{G}^*_{H|B}, 0) < \Theta_L(G_B, 0) \leq \Theta_L(\hat{G}^*_{H|B}, q) \tag{27}$$

for some $G_B \in [0, \hat{G}^*_{H|B}]$.

In other words, we can find parameter values that make some of the conditions for constructing pooling and separating equilibria mutually exclusive which means that we can pick only one of the equilibria in propositions 2 and 3.

3. Conclusion

Third-party intervention in conflicts could lead to perverse outcomes (e.g., Regan, 2002). In this paper, I found an effect dubbed the *indirect* Samaritan's dilemma where a third-party's intervention in a conflict does not cause her beneficiary to engage in undesirable behavior but causes those in strategic relationships with her beneficiary to engage in undesirable behavior. In the present model, the *expectation* of the third party's assistance *energizes* the enemy of her ally in the current period and this behavior is purely driven by strategic considerations. An important implication is that it may be strategically wrong for the third party to base her decision to intervene or withdraw from the conflict on how bad the current situation is.

The indirect Samaritan's dilemma may have applications in other areas. For example, in Bernheim and Severinov (2003) an altruistic parent plays a signaling game with two children who do not strategically interact with each other. It might be interesting to investigate a variation of the model in Bernheim and Severinov (2003) with *sibling rivalry* (e.g., keeping up with the Joneses) and a biased parent who does not know the wealth of her children. The children send signals about their wealth to their parent through their consumption levels which are increasing in wealth and their parent's choice of bequests is based on her beliefs about their wealth.

There are several ways to extend the analysis. One extension is to consider the case where both factions in the conflict have an ally who is willing to offer military assistance. This case could be challenging and could lead to interesting strategic interactions between the third parties that are not explored in this paper especially if the third parties move sequentially. For example, a third-party's assistance to her ally may give some information to the other third party about what the third party knows about her ally’s type and her inference about the type of the ally’s enemy. Another extension is to make the type space continuous.

Appendix A

Proof that the signaling game does not satisfy the single-crossing condition

Consider the signaling game between the third party and faction B. To write faction B's preferences defined over G_B and M, one must do this conditional on what faction A does. Notice that G_B does not enter faction B's payoff function in period 2. Accordingly, conditional on faction B's *de facto* first-mover role, I write his payoff function conditional on (a) faction A's best response behavior for a given G_B in period 1, and (b) the equilibrium of the game between factions A and B in period 2 for a given M.[27]

Using (14) and (17) in the text and some arbitrary valuation, W, we can write faction B's payoff as

$$\Theta = W\sqrt{\frac{G_B}{V}} - G_B + \left(1 - \frac{(1+M)^{\theta}V}{2W_k}\right)\left(W_k - \frac{(1+M)^{\theta}V}{2}\right) \tag{A1}$$

The slope of his indifference curve (iso-payoff curve) is defined as

$$\eta \equiv -\frac{\partial\Theta/\partial G_B}{\partial\Theta/\partial M} \tag{A2}$$

Then

$$\operatorname{sign}\left(\frac{\partial\eta}{\partial W}\right) = \operatorname{sign}\left(V^2(1+M)\sqrt{\frac{G_B}{V}} + W^2(1+M)^{1-\theta} - VW(1+M)\right) \tag{A3}$$

Suppose $V \geq W$. Then (A3) implies that $\partial\eta/\partial W < 0$ as $G_B \to 0$ but $\partial\eta/\partial W > 0$ as $G_B \to \infty$. Therefore, the indifference curves do not satisfy the single-crossing condition.

---

[27]For faction A, we can his payoff function for some arbitrary investments in periods 1 and 2 by faction B and also show that is does not satisfy the single-crossing condition.

Appendix B

Choosing the value of λ in the out-of-equilibrium beliefs specified in (18a)

Note that $\Theta_k(\hat{G}_B,1) > \Theta_k(\hat{G}_B,q)$. Since $\hat{\Pi}_k^1$ is monotonically decreasing in $\hat{G}_B$ for $\hat{G}_B \in [\hat{G}^*_{H|B}, V]$, it follows that $\Theta_k(\hat{G}_B,1)$ is monotonically decreasing in $\hat{G}_B$ for $\hat{G}_B \in [\hat{G}^*_{H|B}, V]$, k = H,L. These two observations imply that there are unique and positive values of λ_k which solve

$$\Theta_k(\hat{G}_B + \lambda_k, 1) = \Theta_k(\hat{G}_B, q), \tag{B1}$$

k = H,L. Denote the solutions by $\hat{\lambda}_k > 0$.

If $\hat{\lambda}_H = \hat{\lambda}_L$, then we set $\lambda = \hat{\lambda}_H = \hat{\lambda}_L$ and we are done. Without loss of generality, suppose $\hat{\lambda}_H > \hat{\lambda}_L$.[28] Then the following statements hold:

If $G_B \in [\hat{G}_B + \hat{\lambda}_L, \hat{G}_B + \hat{\lambda}_H)$, then

$$\Theta_H(G_B,1) > \Theta_H(\hat{G}_B,q) \text{ but } \Theta_L(G_B,1) \le \Theta_L(\hat{G}_B,q). \tag{B2}$$

If $G_B \in [\hat{G}_B, \hat{G}_B + \hat{\lambda}_L)$, then

$$\Theta_k(G_B,1) > \Theta_k(\hat{G}_B,q), \tag{B3}$$

k = H, L.

Given the argument in the text, the statement in (B3) implies that we must set $\lambda = \hat{\lambda}_L$ in (18a). Based on the intuitive criterion, the statement in (B2) implies that the

[28]My results do not hinge on the sign of this inequality. This inequality need not hold because while a given increase in investment is more costly to the weak type in period 1, a decrease in assistance in the form of a subsidy to faction A (as modeled here) due to the increase in the third-party's belief (i.e., μ) from q to 1 is more valuable to the weak type of faction B because for a given reduction in military assistance, faction A reduces his investment in period 2 by a bigger amount if faction B is weak than if faction B is strong.

third party must set $\mu = 1$ for $G_B \in [\hat{G}_B + \hat{\lambda}_L, \hat{G}_B + \hat{\lambda}_H)$.[29] And for $G_B \in [\hat{G}_B + \hat{\lambda}_H, \infty]$, no type has the incentive to deviate from the equilibrium even if the third party believes that he is strong, so $\mu = 1$ is a reasonable belief. Therefore, it is reasonable for the third party to believe that $\mu = 1$ for $G_B \in [\hat{G}_B + \hat{\lambda}_L, \infty]$. All of these arguments result in the beliefs in (18a).

Appendix C

Alternative formulation of third party intervention

In the text, the third-party intervention in period 2 was modeled as a subsidy to faction A's arms expenditure. Suppose instead that the third party is directly involved in the conflict in period 2 and so chooses her investment in arms, $X_C = X_C(\mu)$, where her belief that faction B is strong is $\mu \in [0,1]$. Let the cost of X units of arms be X. Alternatively, the formulation below need not imply that the third-party's soldiers are directly involved in the conflict. It is also consistent with the interpretation that the third party makes transfers to faction A but the transfers are in-kind transfers in non-human military goods as opposed to subsidies that are conditional on faction A's investment in military goods or arms.

As before, in period 2 the third-party moves first, followed by faction A, and then faction B. In this period, faction B cannot signal his type since the third party moves before he does.

Conditional on $X_{k|A}$ and $X_C(\mu)$, faction B of type k chooses $X_{k|B}$ to maximize

[29] If $\hat{\lambda}_H < \hat{\lambda}_L$, then according to the "intuitive criterion" we should set $\mu = 0$ over this range of investments. This does not change any of the results.

$$\Pi^2_{k|B} = \frac{X_{k|B}}{X_{k|B} + X_{k|A} + X_C} W_k - X_{k|B}. \tag{C1}$$

This gives her best-response function in period 2 as

$$X_{k|B} = \sqrt{W_k (X_{k|A} + X_C)} - X_{k|A} - X_C, \; k = H, L.$$

Conditional on $X_C(\mu)$, faction A chooses $X_{k|A}$ to maximize

$$\Pi^2_{k|A} = \frac{X_{k|A} + X_C}{X_{k|B} + X_{k|A} + X_C} V - X_{k|A}. \tag{C2}$$

It is clear from (C2) that the contributions of faction A and the third party are voluntary contributions to a public good.

Putting $X_{k|B} = \sqrt{W_k (X_{k|A} + X_C)} - X_{k|A} - X_C$ into faction A's payoff function and maximizing gives faction A's investment in period 2 when his opponent is faction B of type k as

$$X_{k|A} = \frac{V^2}{4W_k} - X_C, \tag{C3}$$

k = H, L.

Note from (C3) that $X_{k|A} + X_C = 0.25V^2 / W_k$. For a given W_k, the optimal aggregate investment in arms by the third party and faction A is unique. Given that the third party moves first, she will optimally choose $X_C = 0$ knowing that faction A will choose $X_{k|A} = 0.25V^2 / W_k$. Therefore, the third party will never intervene. If S = V and the third party and faction A move simultaneously, then the optimal aggregate investment is unique but the individual investments are not.[30] The party with the higher valuation contributes the entire investment in arms if $V \neq S$. This result hinges on the constant marginal cost of arms (see Konrad, 2009). Bergstrom, Blume, and Varian (1986) obtain a similar result in the general case of the voluntary contribution to public goods.

A model of third party intervention in which the third party never intervenes is not a good model. We can rectify this by allowing the third party to move after faction A has moved but before faction B moves. In that case, faction A will not invest in period 2 and will leave the entire cost of fighting faction B to the third party. This strikes me as unsatisfactory.

Suppose we maintain the original timing of moves. Being allies, we would expect faction A and the third party to negotiate an agreement on what to do and not fully free ride on each other's investment. For example, the third party, who presumably has more resources, might make his investment decision conditional on faction A investing some minimum amount $\overline{X}_A > 0$ in arms. Alternatively, given that I have used the expressions "armed investment" or simply "arms" to represent a composite military good made up of soldiers and weapons, the third party could ask faction A to provide $\overline{X}_A > 0$ soldiers while she provides the weapons for the production of the composite military good. Assume that faction A honors this agreement by investing in this minimum and fixed amount.

Conditional on $X_{k|A} = \overline{X}_A$, the third party chooses X_C to maximize

$$\Pi_C^2 = \mu \frac{\overline{X}_A + X_C}{X_{H|B} + \overline{X}_A + X_C} S + (1-\mu) \frac{\overline{X}_A + X_C}{X_{L|B} + \overline{X}_A + X_C} S - X_C . \tag{C4}$$

[30] For increasing and strictly convex cost functions, the individual contributions will be unique (Esteban and Ray, 2001).

Putting $X_{k|B} = \sqrt{W_k(\overline{X}_A + X_C)} - \overline{X}_A - X_C$ into (C4) and maximizing gives

$$X_C^* = \frac{1}{4}\left(\frac{\mu}{\sqrt{W_H}} + \frac{1-\mu}{\sqrt{W_L}}\right)S^2 - \overline{X}_A \tag{C5}$$

I assume that S is sufficiently high so that $X_C^* > 0$. Then given $W_H > W_L > 0$, (C5) implies that $\partial X_C^* / \partial\mu < 0$. Then applying the envelope theorem to faction B's payoff function in (C1), it follows that faction B's equilibrium payoff in period 2 is increasing in μ. Hence the propositions in the text continue to hold under this alternative formulation of third-party intervention.

References

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

Andolfatto, D. (2002). A Theory of Inalienable Property Rights. *Journal of Political Economy* 110: 382-393.

Baik, K.H., and Shogren, J.F. (1992). Strategic Behavior in Contests – Comment. *American Economic Review* 82: 359-362.

Baliga, S., and Sjöström, T. (2008). Strategic Ambiguity and Arms Proliferation. *Journal of Political Economy* 116: 1023-1057.

Banks, J.S., and Sobel, J. (1987). Equilibrium Selection in Signaling Games. *Econometrica* 55: 647-661.

Bergstrom, T., Blume, L., Varian, H.R. (1986). On the Private Provision of Public Goods. *Journal of Public Economics* 29: 25-49.

Bernheim, B. D., and Andreoni, J. (2009). Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects. *Econometrica*, forthcoming.

Bernheim, B.D., and Severinov, S. (2003). Bequests as Signals: An Application for the Equal Division Puzzle. *Journal of Political Economy* 111: 733-764.

Bernheim, B.D., and Stark, O. (1988). Altruism within the Family Reconsidered: do Nice Guys Finish Last? *American Economic Review* 78: 1034-1045.

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. (2009). Civil War. *Journal of Economic Literature*, forthcoming.

Blouin, M., and Pallage, S. (2009). Addressing the Food Aid Curse. *Economics Letters* 104: 49-51.

Bruce, N., and Waldman, M. (1990). The Rotten-Kid Theorem Meets the Samaritan's Dilemma. *Quarterly Journal of Economics* 105: 155-165.

Buchanan, J.M. (1975). The Samaritan's Dilemma. In *Altruism, Morality, and Economic Theory*, Edmund S. Phelps (Ed), New York: Russell Sage Foundation, 71-85.

Carrillo, J.D., and Mariotti, T. (2000). Strategic Ignorance as a Self-Disciplining Device. *Review of Economic Studies* 67: 529-544.

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.

Cho, I-K., and Sobel, J. (1990). Strategic Stability and Uniqueness in Signaling Games. *Journal of Economic Theory* 50: 381-413.

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

Coate, S. (1995). Altruism, the Samaritan's Dilemma, and Government Transfer Policy. *American Economic Review* 85: 46-57.

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

Dixit, A.K. (1987). Strategic Behavior in Contests. *American Economic Review* 77: 891-898.

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

Esteban, J., and Ray, D. (2001). Collective Action and Group Size Paradox. *American Political Science Review* 95: 663-672.

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

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

Grossman, S., and Perry, M. (1986). Perfect Sequential Equilibria. *Journal of Economic Theory* 39: 97-119.

Horner, J., and Sahuguet, N. (2007). Costly Signaling in Auctions. *Review of Economic Studies* 74: 173-206.

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

Lagerlof, J. (2004). Efficiency-Enhancing Signaling in the Samaritan's Dilemma. *Economic Journal* 114: 55-68.

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

Lindbeck, A., and Weibull, J.W. (1988). Altruism and Time Inconsistency: the Economics of Fait Accompli. *Journal of Political Economy* 96: 1165-1182.

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

Morgan, J., and Vardy, F. (2007). The Value of Commitment in Contests and Tournaments when Observation is Costly. *Games and Economic Behavior* 60: 326-338.

Munster, J. (2009). Repeated Contests with Asymmetric Information. *Journal of Public Economic Theory* 11: 89-118.

Pedersen, K.R. (2001). The Samaritan's Dilemma and the Effectiveness of Development Aid. *International Tax and Public Finance* 8: 693-703.

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.

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

Slantchev, B.L. (2009). Feigning Weakness. *International Organization*, forthcoming.

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

Zhang, J., and Wang, R. (2009). The Role of Information Revelation in Elimination Contests. *Economic Journal* 119: 613-641.