Interventions and Conflict Incentives

Recent literature suggests that well-intentioned third party intervention in military conflict can lead to moral hazard by acting as a subsidy to rebellion. In this paper, I suggest (i) that failure to intervene can also lead to conflict, which I call the “insufficient deterrence” problem, and (ii) a classification scheme for potential conflicts in the world, which can be used to derive the optimal intervention policy on case-by-case basis. The optimal policy may be to intervene at random since certainty in some cases will encourage either parties who benefit from intervention or parties who lose to attack. The model also suggests that post-conflict aid and a siding-with-the-winner approach may promote the incidence of conflict, even if such policies lower the loss from unavoidable conflicts.
Thorsten Janus  
University of Wyoming  
College of Business  
Department of Economics and Finance, Dept 3985  
1000 E. University Ave.  
Laramie, WY 82071  
USA  
tjanus@uwyo.edu  
Tel. 307-760-8309  
Fax 307-766-5090
1. **Introduction**

Third party intervention in internal or between-polity conflicts has a history almost as long as warfare itself. Among the driving forces, external parties may have direct economic or political stakes, sympathize with one of the parties to conflict, suffer spillovers, or be concerned with humanitarian issues. Globalization, resource scarcity, and state weakness in many countries make the current upward trend in the number of interventions unlikely to reverse in the near future.

Most of the research on intervention has focused on the normative issue of how and when interventions should occur to end conflict in a rapid and painless way. The underlying concern is that conflict slows economic, social, political, and human development and should therefore be avoided. However, while intuitive and morally appealing, recent work suggests an important qualification: if the parties to conflict are forward looking and know that intervention can help their situation in case of conflict, then intervention *ex-post* can promote conflict *ex-ante*. This so called “moral hazard” problem with interventions is discussed at length in a recent special issue of *Ethnopolitics* (Volume 4, Issue 2, 2005).¹ The risk of moral hazard implies that taking conflict as a given when analyzing the effect of interventions can be a mistake, and that the traditional view that an efficient benevolent intervention is always welfare improving, should be reversed in some cases: sometimes less willingness or ability to intervene is preferable. By extension, the threat of intervention by a self-interested “fourth” party who will destitute the groups in conflict may be preferable.
However, while many cases seem to support the moral hazard view (Kuperman 2005), not only will the absence of intervention be insufficient to prevent all conflicts, but also there may also be cases where precisely the absence of intervention will lead to conflict. This leaves the optimal design of intervention policy for a benevolent international coalition unclear: should it intervene always, never, or under which circumstances? This paper aims to answer that question by proposing a classification scheme for potential conflicts in the world, and using the scheme to derive a procedure for finding the optimal intervention policy, given the parameters of each particular conflict.

To derive the procedure for finding case-by-case optimal intervention policies I use, first, the concept of moral hazard. I will say that moral hazard (in the context of humanitarian interventions) occurs when at least one party to a potential conflict assigns a positive likelihood to intervention, should a conflict occur, and this expectation is sufficient to cause the conflict. Second, I introduce the “opposite” concept of insufficient deterrence. Insufficient deterrence occurs when at least one party to a potential conflict assigns a positive likelihood to no intervention, should a conflict occur, and this expectation is sufficient to cause the conflict.

In addition, a general normative analysis of interventions must look not just at whether conflict occurs, but at the total welfare costs of different intervention policies. In particular, although in some cases the right intervention decision can prevent the conflict, in other cases the conflict is unavoidable and the intervention must be used to maximize
global welfare given the occurrence of the conflict. The paper considers these cases as well.

Finally, my world welfare criterion will simply be the global resource cost of conflict. This may seem limited because it is a narrow economic measure and is unresponsive to distributional and fairness concerns. For example, if one rich party simply appropriates the wealth of another, poor party, without provocation and without battle, there will be no decline in welfare. However, in defense of the measure, available resources can be broadly interpreted as the monetary value of the economic, political and social resources, or rights and obligations, available to all groups jointly. Group wealth shares can be similarly interpreted. In addition, even a narrow focus on economic value is common in the literature (see the literature review below), distributional and fairness assessments are subjective and controversial, and the focus of the moral hazard analyses the paper seeks to relate to is on minimizing the direct losses from conflict – without concern for status quo injustices. Finally, a traditional defense for focusing on aggregate wealth maximization in economics (which inspired the moral hazard argument) is that, with more wealth, there is enough to make at least one party better off and no party worse off. However, in the end, in an imperfect world, certain wars may simply be worth fighting (for example, the alternative in WWII was to appease the Axis powers). Therefore, using intervention policy to prevent all wars, even were it feasible, is not obviously the right thing to do either.
The rest of the paper is organized as follows. First, section 2 reviews the most closely related literature. Section 3 then presents a formal model and derives a classification scheme for potential conflicts in the world, which can be used to design the optimal intervention policy on a case-by-case basis. I also discuss the implications of post-conflict aid and a siding-with-the-winner approach for the likelihood of moral hazard and insufficient deterrence problems. Section 4 concludes the paper by summarizing and discussing some caveats.

2. Related Literature

In the literature discussing the incentive effects of interventions, Carment and Rowlands (2001) analyze a Nash equilibrium model with two parties to a conflict. They show that the effect of third party intervention on individual and total fighting efforts depends on the bias and capability of the third party. However, because Carment and Rowlands only consider conflicts already begun, they ignore the incentives to begin a conflict in the first place which are main concern of this paper. They also consider a single third party only, while I allow for “fourth” party to enter if the benevolent third party decides not to enter the conflict. Carment and Rowlands also do not consider commitment issues for the third party, partly because they take the conflict itself as a given. The same is true for the papers by Chang et al., (2007) and Amegashie and Kutsoaki (2007).

Findley and Teo (2006) present an informal model of intervention by several outside parties, where the decisions of each party depends on the conflict and the decisions of the
other, as in the case of the superpowers in the Third World during the cold war. Their paper, while closely related, differs from mine in analyzing intervention by two self-interested third parties and having a positive focus. In contrast, this paper has a normative focus and is focused on the decision problem of an outside party acting to minimize the global resource losses from conflict. Findley and Teo’s model may be particularly relevant for understanding Cold War interventions and conflicts where UN and other “benevolent” outsiders are unlikely to be effective, while this paper is more relevant for the Post Cold War era and the design of future international cooperative efforts to promote peace.

Crawford (2005) and Kuperman (2005) argue that rebels may promote genocidal retaliation against the group they claim to represent, since this would trigger humanitarian intervention from which they benefit at the expense of the government. In other words, rebels trigger violence against their own people to provoke sympathetic interventions. This is the moral hazard problem discussed above, since without the hope of intervention the rebels would not trigger the violence. However, the moral hazard argument assumes that the intervener mistakenly sympathizes with the rebels who, after all, cause the suffering, that the government is willing to engage in genocidal violence which will hurt it, and that the victims of genocidal violence do not turn against the rebels. The rebels must therefore outsmart both the government, international society, and the civilians they claim to fight for. While these assumptions may be justifiable in some cases, the actors in this paper are fully rational and informed, and they are not outsmarted. Both Crawford and Kuperman also focus only on intervention biased in favor
of a weak rebel group, while the model below treats groups symmetrically with any one

group arbitrarily strong, allows both neutral and biased interventions, allows more than

one intervener, and considers the role of commitment. Most importantly, neither

Crawford nor Kuperman consider the incentive effects of not intervening.

Finally, the papers by Crawford and Kuperman are part of the special issue of

*Ethnopolitics* devoted to the moral hazard problem. However, because of the shared

starting point, the comments made above tend to apply to the other papers in the issue as

well. On the other hand, those papers in the issue which criticize the moral hazard

argument offer no modified version or a modification which only partly addresses the

limitations of the argument. None of the other papers consider the insufficient deterrence

problem. Nonetheless, all of these papers are fundamentally concerned with the same

dependence of conflict incentives on interventions as here.

### 3. The Incentive Effects of Interventions and Optimal Policy

Suppose a situation where group 1 can choose to attack (action a) or not attack (na) group

2. One of these may be a government. If group 1 attacks, a multinational peace enforcing

M, can choose to intervene (i) or not (ni). For convenience, one can take M to represent

the United Nations. M could also be a peacemaking or peacekeeping force, but then both

parties must be better off with the intervention since by conventional definitions these

operations require the consent of all parties. I discuss this case below, but for now assume

a peace enforcing M.³
The value of initial resources (wealth) available in society is $\theta$. However, conflict followed by intervention or conflict without intervention reduces the wealth to $\theta - x$ or, respectively, $\theta - y$, due to negative social, economic and other effects of conflict on society. $s_1^i$ ($s_2^i$) is the, perhaps subjectively, expected wealth share of group 1 (group 2) if intervention occurs and the notation is similar for no intervention and no attack by group 1. Consistent with reality, M cannot change the wealth shares at will, but only affect them insofar as they happen to depend on M’s intervention, which may change the duration or course of the conflict.

I will initially assume that group 2 is passive in the sense that it does not initiate conflict. Allowing group 2 to begin the conflict as well will modify the results in some cases, however, and so the current passivity assumption is made to clarify the exposition and not merely for simplicity. However, it should become clear why the approach is sensible.

The extensive form of the game between the groups and the intervener is depicted in the game tree in figure 1, where the second payoff at each terminal node is the payoff to group 2 and the third payoff is M’s. M’s payoff is the global wealth loss due to conflict, which is the sum of the cost incurred by the parties to conflict and the direct cost of the intervention $c_M$. I use the convention that indifference between two decisions for any party leads to no action. Appendix I considers why Coasean bribes (Coase, 1960) can fail to prevent conflict. I will assume, unless otherwise stated, that M cannot commit to
intervening or not intervening \textit{ex-ante}. It therefore intervenes if and only if intervention lowers the world resource loss from the \textit{ex-post} perspective.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{diagram.png}
\caption{The Intervention Game}
\end{figure}

\textbf{Solution of the game}

The game can be solved by backward induction. First, if an attack by group 1 has occurred, M intervenes if an only if

\begin{equation}
 x + c_M < y
\end{equation}

which says that the global wealth loss from conflict with intervention is less than the loss without intervention. Second, if group 1 expects intervention, it attacks if and only if

\begin{equation}
 s_i^i(\theta - x) > s_i^{na}\theta.
\end{equation}
that is, group 1’s payoff is higher with conflict followed by intervention than without conflict. On the other hand, if group 1 expects no intervention, it attacks if and only if

\[ s^{ni}_i (\theta - y) > s^{na}_1 \theta , \]  

that is, group’s payoff is higher under conflict without intervention than without conflict.

If group 1 attacks solely when it expects intervention, there is *moral hazard*. This requires that (1) holds – otherwise M will not intervene - and that

\[ s^1_i (\theta - x) > s^{na}_1 \theta \geq s^{ni}_i (\theta - y) . \]  

When there is moral hazard, since intervention prompts conflict followed by intervention, there is a global wealth loss of \( x + c_M > 0 \).

Conversely, however, group 1 may attack solely because it expects *no* intervention. This applies if (1) does not hold – so that M will prefer not to intervene if there is conflict - and if the payoff ranking for group 1 is

\[ s^{ni}_i (\theta - y) > s^{na}_1 \theta \geq s^i_1 (\theta - x) , \]  

which implies *insufficient deterrence*, that is, the expectation of no intervention is the cause of conflict. The expected failure to intervene leads to a global loss of wealth equal to \( y > 0 \).
Moral hazard is more likely to occur ((4) is more likely to hold) if \( y - x \) is large, so that intervention substantially lowers the resource loss from conflict, and if \( s_i^i - s_i^{ni} \) is large, that is, if the attacking group 1 has a substantial wealth share gain from intervention compared to no intervention. The wealth share of group 1 may increase with intervention if otherwise, during the “natural” (intervention-free) course of the conflict its relative position would deteriorate over time. In this case, intervention freezes the wealth distribution at a favorable time for group 1.

Thus, to avoid moral hazard requires either a combination of low values of \( s_i^i - s_i^{ni} \) and \( y - x \), so that (4) does not hold; or, alternatively, a high resource cost of intervention \( x + c_M \), so that (1) does not hold - if (1) does not hold, then M can credibly commit not to intervene, because the global resource cost of intervention exceeds the gain; or a predetermined commitment by M to not intervene. This result puts often criticized features of UN interventions, such as blocking by political opposition, high-cost, and ineffective, in a different perspective: the knowledge that intervention is unlikely and ineffective if it occurs may help to prevent the moral hazard problem.

Conversely, the insufficient deterrence problem expressed by (5) is more likely when the opposite parameter configurations apply: when \( x - y \) is large, so that intervention fails to lower the wealth loss, or when \( s_i^{ni} - s_i^i \) is large, so that intervention greatly reduces the wealth share of group 1. To avoid the insufficient deterrence problem requires either a
combination of a low values of $s_i^m - s_i^l$ and $x - y$, so that (5) does not hold; or a low resource cost of intervention $x + c_M$, so that (1) holds – then $M$ can credibly commit to intervene ex-post; or a commitment by $M$ to intervene.

Combining these analyses of the moral hazard and insufficient deterrence problems suggests that, as long as there are some possible conflict cases in the world where intervention would lead to conflict, and some cases where the absence of intervention would do so, there can be no one-size-fits-all approach to interventions. Instead, intervention policy, including commitments, should be determined on a case-by-case basis. To illustrate, those who subscribe to the Save Darfur movement suggests that the UN may have been able to prevent the genocide by committing to intervention if genocide were to occur – perhaps by establishing a rapid and powerful intervention force with a strong mandate. This would have conveyed to the Sudanese governments that “we will intervene if you attack”. Alternatively, if a moral hazard view of the conflict in Darfur is appropriate, the UN could have prevented the genocide by committing *not* to intervene, or by ensuring that intervention would be costly and ineffective - due, for example, to a weak mandate and poor preparation. This would have conveyed to the rebels that “we will not intervene if you attack the government (or the Arab nomads)”.

Paradoxically, however, as I show below, under some conditions the flipside of preventing one side’s aggression may be to encourage that of the other.

Finally, apart from the moral hazard and insufficient deterrence situations expressed by (4) and (5), two other configurations of parameters are possible. First, if
\[
\min \{s_i^1(\theta - x), s_i^m(\theta - y)\} > s_i^{ma} \theta
\]  \hspace{1cm} (6)

then there is neither a moral hazard nor an insufficient deterrence problem: conflict will occur regardless of expected intervention, since group 1 always has a higher payoff from conflict. For such conflicts, the best M can do is to intervene \textit{ex-post} if and only if intervention lowers the global loss from conflict, that is, if (1) holds. Second, if

\[
s_i^{ma} \theta \geq \max \{s_i^1(\theta - x), s_i^m(\theta - y)\}
\]  \hspace{1cm} (7)

then group 1 has a higher payoff from peace than from conflict regardless of intervention. In this case, there is no conflict, no loss, and again neither a moral hazard- nor an insufficient deterrence problem.

\textbf{A weak mandate}

Suppose that M is peacekeeping or peacemaking instead of peace enforcing. Then all parties must agree to intervention for intervention to be possible. I will call this a weak mandate, while peace enforcement is a strong mandate. Formally, intervention with a weak mandate requires

\[
s_i^1(\theta - x) \geq s_i^m(\theta - y)
\]  \hspace{1cm} (8)

and
so that no party will block intervention. (8) and (9) state that groups 1 and 2, respectively, has a higher payoff from conflict with intervention than from conflict without intervention.

How does a weak mandate affect the likelihood of moral hazard and insufficient deterrence? First, if condition (6) holds, there is conflict regardless of the expected intervention decision. However, if additionally either (8) or (9) fail to hold, then one party is better off if there is no intervention and will block it. In turn, as long as (1) holds, there is a resource loss due to the weak mandate of $y - x - c_M > 0$. Thus, while M does not cause the conflict, it fails to minimize the loss from an inevitable conflict. Second, if condition (7) holds, there is no conflict regardless of M’s behavior and no effect of a weak mandate.

Third, if the moral hazard condition (4) holds, then clearly condition (8) holds also: group 1 will not block intervention since intervention is the whole purpose of starting the conflict. However, (9) may not hold and then group 2 will block intervention. This could apply if the resource share of group 2 is rising over time in the natural course of the conflict. If group 2 will block the intervention group 1 is hoping for, the weak mandate will prevent moral hazard and the conflict will not occur. Intuitively, making it easier to block interventions can prevent moral hazard because, by definition, the conflict occurs
precisely due to the anticipated intervention. In the Darfur example, if we suppose that it is easy for the Sudanese government to block intervention, the rebels cannot reasonably be hoping that intervention will help them at the expense of the regime. This would make the moral hazard explanation of the conflict less convincing. However, in reality the regime does appear at least somewhat susceptible to international pressure.

Finally, if the insufficient deterrence condition (5) holds, then clearly (8) does not hold: group 1 starts the conflict precisely because it expects no intervention and will therefore block any proposed intervention. A weak mandate therefore promotes the likelihood of insufficient deterrence. Returning to Darfur, if to begin with the regime believed that it could easily block outside interference once the genocide began, and that belief was critical to beginning the genocide, this would support the view of the Save Darfur movement that international passivity is to blame.

Allowing group 2 to be active

I now finally allow group 2 to be active. This involves more than just label switching between groups, because, as long as expectations are realistic, the wealth shares across groups must sum to one (in the next section, I discuss the implications of overconfidence). Thus, in the two group case considered so far, with realistic expectations, we must have $s^j_2 = 1 - s^j_i$, $j = i, ni, na$. Importantly, as I show formally just below, this means that the flipside of a moral hazard facing group 1 may be insufficient deterrence facing group 2. In other words, if group 1 wants to fight if and only if there is
intervention, then group 2 may want to fight if and only if there is no intervention, suggesting that conflict could be inevitable. This could apply to Darfur if, for example, an early international commitment to intervene would have caused rebel attacks but a commitment not to intervene would have caused government attacks – a choice between two evils. Formally, using $s_{2}^{i} = 1 - s_{1}^{i}$, $j = i, ni, na$, the first inequality in the moral hazard condition (4) implies

$$s_{2}^{na} \theta > s_{2}^{i}(\theta - x) + x$$

while the second inequality in (4) implies

$$s_{2}^{i}(\theta - x) + y \geq s_{2}^{na} \theta.$$  

Combining these inequalities, we get

$$s_{2}^{ni}(\theta - y) + y \geq s_{2}^{na} \theta > s_{2}^{i}(\theta - x) + x$$  \hspace{1cm} (10)

and for $x$ and $y$ sufficiently small, the insufficient deterrence condition for group 2,

$$s_{2}^{ni}(\theta - y) > s_{2}^{na} \theta \geq s_{2}^{i}(\theta - x)$$  \hspace{1cm} (11)

will hold. Thus, as claimed, group 2 faces insufficient deterrence precisely when group 1 faces moral hazard. The former problem calls for intervention, but the latter calls for no
intervention, so not even a commitment to intervene or not will prevent conflict. Symmetrically, group 1 facing insufficient deterrence implies that group 2 faces moral hazard when $x$ and $y$ are small. The result that, for $x$ and $y$ small, the flipside of moral hazard for one group is insufficient deterrence for another group is sensitive to the number of groups. However, I show in appendix II that moral hazard always, for $x$ and $y$ sufficiently small, implies either insufficient deterrence for at least one other group or that at least one other group prefers conflict regardless of intervention. Thus, moral hazard, for $x$ and $y$ sufficiently small, always makes conflict inevitable; commitment to either intervene or not intervene does not matter. By symmetry, insufficient deterrence always, for $x$ and $y$ sufficiently small, imply either moral hazard for at least one other group or that at least one other group prefers conflict regardless of intervention.

Does this mean that interventions can only be used to minimize the resource losses from conflict and never to prevent them – that the entire moral hazard and insufficient deterrence debate is pointless? The answer is no for at least two reasons:

(i) First, the conflict losses $x$ and/or $y$ may be sufficiently large that only one group wants to fight. This is true for any number of groups, but to be concrete I will frame the discussion in terms of two groups. Large values of $x$ or $y$ gives rise to two types of conflicts which could be avoided by proper choice of intervention policy. First, for $y$ large, (11) becomes the equivalent of (7) for group 2, and so the only remaining problem is a moral hazard problem facing group 1. Therefore, if M can only resolve this one
single incentive problem, the conflict can be prevented. Call this avoidable conflict type A. Type A avoidable conflicts may more precisely capture the moral hazard problem than we did above: in type A conflicts, rebels attack because they expect intervention and, at the same time, the government is uninterested in conflict with or without intervention. Therefore, an international commitment not to intervene would maximize global welfare, saving both the direct cost of conflict and the cost of intervention. Examples of Type A avoidable conflicts may be the Bosnia and Kosovo in the 1990s, as argued by Kuperman (2005); Chechnya, if international pressure helped decrease Russia’s punishment for irredentism; aggression by minority factions in Iraq, who were protected by the coalition forces; and rebellion in Aceh province in Indonesia if it was due to international pressure on the Indonesian government. Generally, many conflicts today are begun by rebels against their states and international society often favors weak and non-state parties, consistent with the moral hazard view.

Second, for large, (11) still holds, but (4) becomes (7) – in other words, group 1 no longer gains from conflict either with or without intervention and the only remaining problem is the insufficient deterrence problem facing group 2. Therefore, if M can only resolve that one single incentive problem, conflict can be avoided. Call this avoidable conflict type B. In a type B conflict, one party attacks because it expects no intervention and the other party prefers not to attack with or without intervention. When parameters make for a conflict type B, and the groups are a government and a (potential) rebel group, the moral hazard problem is reversed: the government (and not the rebels) attacks because it expects no intervention and, at the same time, the potential rebels are
uninterested in conflict with or without intervention. Therefore, an international commitment to intervene in case of conflict would prevent conflict and maximize global welfare. Examples of type B avoidable conflicts may include Rwanda’s genocide in 1994, Turkey’s contemporary conflict with resident Kurds, Germany’s persecutions and annexations during World War II, and Turkey’s Armenian genocide in the early 20th Century. Also, at least popular sentiment in developed countries blames the regime, rather than the rebels, for the genocidal violence in Darfur. Generally, while the sovereignty principle may have weakened in recent years, many countries still have governments who can pose a substantial threat to, and are not representative of, minorities, and who can prevent foreign intervention. Obviously, however, rather than claim categorically that there are too many or too few interventions in the world (too many type A/moral hazard conflicts or too many type B/insufficient deterrence conflicts), it is more reasonable to go case-by-case.

(ii) Second, if group 1 does face moral hazard and group 2 insufficient deterrence – if (4) and (11) hold simultaneously - then we can think of the problem being the following: the certainty of intervention means that group 1 can ignore the loss if there is no intervention, while the certainty of no intervention allows group 2 to ignore its loss in case of intervention. One may guess, therefore, that if M can convince the groups that it may or may not intervene, then the scope for loss can potentially discourage both parties. Formally, suppose now that M intervenes with probability $\tilde{p}$. To discourage both groups from fighting, this probability must satisfy
which state that the expected payoff from conflict is less than the payoff without conflict for each group, and where as above I assume realistic expectations, so that $s_2^i = 1 - s_1^i$.

Suppose now that we pick $\tilde{p}$ to solve (12a) with equality. Then substitution into (12b) gives

$$\tilde{p}x + (1 - \tilde{p})y \geq 0$$

which is always true for $\tilde{p}$ in $[0,1]$. $\tilde{p}$ in $[0,1]$ solving (12a) always exists, by the intermediate value theorem, since in (4) the intervention payoff $s_1^i(\theta - x)$ and the no intervention payoff $s_1^{ni}(\theta - y)$ are on opposite sides of the no-conflict payoff $s_1^{na}\theta$, and (12a) is a continuous function of $\tilde{p}$. Thus, an appropriate randomness in intervention can discourage both the moral hazard problem and the insufficient deterrence problem which may is its flipside for small $x$ and $y$. This result is consistent with reality to the extent that it is usually hard to predict where, when, and how UN and other multilateral interventions will occur.
Kuperman (2008b) argues that randomized intervention policy may be insufficient to
deter conflicts in practice because interventions are already very unpredictable and the
parties to conflict may be able to influence the intervention decision through political
channels. However, first, what is needed is not merely unpredictable intervention, but
sufficiently rare intervention – the right value of $p$. Second, though international
interveners may be subject to pressure from the groups in conflict, this is not obviously
different from a government being subject to pressure from domestic interest groups –
and such problems can, and have, been at least partly resolved by appropriate institutional
design or the concerns of policymakers to preserve a reputation for looking out for the
general interest. The institutions organizing interventions could be similarly constrained.
In addition, it may be in the long-term interest of a country to commit itself to not disturb
the randomization procedure of the intervener if ever the country were to face risk of civil
conflict. For places such as the Balkans, which may be prone to future conflicts and
likely to exert a great pull from interveners due to factors such as history and location,
special institutional arrangements could be used to make future interventions there
particularly difficult. Nonetheless, the independence of policies from affected interests at
any point in time is an empirical question and is certainly important to keep in mind in
every situation of potential conflict.

Are expectations realistic?

In the section above I assumed realistic expectations regarding wealth shares and
aggregate losses from conflict. This was used to derive the result that, for $x$ and $y$ small,
moral hazard and insufficient deterrence may be flipsides of each other. I think that realistic expectations is a reasonable assumption in most cases and evidence that parties are overconfident during conflict does not mean that they were overconfident when they decided to begin the conflict. However, if expectations are indeed unrealistic, they are probably biased towards overconfidence about the conflict payoffs. Such overconfidence makes conflict more likely and narrows the scope for the right intervention policy to prevent conflict, as can be seen formally from (4)-(7). Intuitively, if conflict comes with optimism, expected conflict payoffs are more likely to exceed payoffs without conflict. Thus, overconfidence about conflict payoffs, with or without intervention, will tend to turn the moral hazard condition (4), the insufficient deterrence condition (5), and the peace-regardless-of-intervention condition (7) into the war-regardless-of-intervention condition (6).

While overconfidence about the gains to conflict tends to raise the incidence of conflict, however, it need not raise the incidence of either moral hazard or insufficient deterrence. For example, with very high overconfidence all conditions like (4) and (5) in societies around the world would come to look like (6). There would then be more conflicts, but fewer would be interveners’ fault. On the other hand, suppose that without overconfidence condition (7) holds in all societies – nobody prefers conflict, regardless of intervention. Since the two conflict payoffs on the right hand side of (7) are generally different, if the overestimate is smoothly and strictly increasing in the realistic estimate, moral hazard and insufficient deterrence situations could appear. For example, suppose that for some party the payoffs to peace and to conflict with and without intervention,
respectively, are 5, 4 and 3. Then (7) holds: \(5 \geq \max\{4,3\}\). However, with 30% overestimation of conflict payoffs, the payoffs to conflict with intervention, peace, and conflict without intervention are ranked \(5.2 > 5 > 3.6\). The moral hazard condition (4) now applies to the party, and if it expects intervention, it will attack. Thus, while overconfidence about conflict payoffs promotes conflict incidence, whether it promotes moral hazard or insufficient deterrence is an empirical question. However, to gauge overconfidence, and to determine the payoff rankings of different parties generally, it may be a poor idea to ask the parties themselves: their reporting incentives may be biased to favor certain intervention policies. Kuperman (2008a) considers a range of other types and sources of divergent expectations across parties to conflict.

**Many groups**

I now consider a general number of groups in conflict. The results of this section also apply to two groups with disobedient splinter factions, unless one can justify why it matters that the groups were once together. With many groups, a condition analogous to (12a) applies to each, with wealth shares \(s_g^j, j = i, ni, na, g = 1,2,3\ldots\). In this case, however, we cannot be sure that \(\tilde{p}\) in \([0,1]\) solving all inequalities exists. For example, suppose that there are three groups, but group 3 gets no wealth share in case of conflict, while group 2 has no wealth share without conflict. Then the left hand side of (12b) is unchanged, but the right hand side is zero, and (12b) is unlikely to hold; group 2 has nothing to lose. Thus, using random interventions to prevent conflict, as described above, may not be possible. However, it is still possible to build intervention policy on the threat
of large losses from conflict: if \( x \) and/or \( y \) are sufficiently large, this again will remove any moral hazard problem, insufficient deterrence problem, or both. M can then deal with any remaining insufficient deterrence (moral hazard) problem, without causing a moral hazard (respectively, insufficient deterrence) problem.

Regarding the effects of a weak mandate, with many groups, a weak mandate matters only if at least one group prefers no intervention once a conflict has begun:

\[
s_g^w(\theta - y) > s_g^i(\theta - x)
\]

for some group \( g = 1, 2, 3 \ldots \). Condition (14) says that group \( g \)'s payoff from conflict without intervention exceeds its payoff from conflict with intervention. This is consistent with one group facing insufficient deterrence, as expressed in (5), but does not imply it. Therefore, if (14) holds for at least one group, then a weak mandate can prevent a moral hazard problem without necessarily causing an insufficient deterrence problem (by permitting aggressors to block intervention).

**Optimal intervention policy**

The analysis above leads to the following procedure to derive the optimal intervention policy on a case-by-case basis. When one of the conditions (4), (5), (6), or (7) – the moral hazard condition, the insufficient deterrence condition, the conflict-regardless-of-
intervention condition and the peace-regardless-of-intervention condition - hold for a group in a potential conflict, I label this case 1, 2, 3, and 4 payoffs, respectively.

1. Identify all potential conflicts situations in the world.

2. For each situation, estimate the wealth of each group \( k = 1, 2, \ldots \) resulting from no conflict \( (s^m_k \theta \text{ in the model}); \) conflict with intervention, \( (s^i_k (\theta - x)) \); and conflict without intervention, \( (s^i_k (\theta - y)) \).

3. For each situation, rank the payoffs for each group and decide if they are case 1, 2, 3, or case 4 payoffs.

4. Ignore situations where payoffs for all groups are case 4 - there will then be peace.

5. In conflicts involving case 3 payoffs for at least one group, there will be conflict regardless of intervention (unless, beyond the scope of the model, interveners can control perceived group payoffs). For these conflicts, intervene if and only if the global resource loss from intervention is less than the loss without intervention, that is, if condition (1) is satisfied.

6. In remaining potential conflict situations, if the payoffs for at least one group are case 1 and no group has payoffs case 2 – there is a moral hazard problem, but no insufficient deterrence problem - then commit, if possible, to not intervene. If the payoffs for at least one group are case 2 and no group has case 1 payoffs– there is an insufficient deterrence problem, but no moral hazard problem - then commit, if possible, to intervene. If both case 1 and case 2 payoffs appear, then commit to an intervention probability satisfying (12a) and (12b) and their possible extensions to
more than two groups. If no such probability exists, then intervene if and only if conflict occurs and the global resource loss from intervention is less than the loss without intervention, that is, condition (1) is satisfied.

7. If M cannot commit in the previous step then, for the remaining potential conflict situations, choose a weak mandate if at least one group has case 1 payoffs (there is a moral hazard problem), no group has case 2 payoffs (there is no insufficient deterrence problem), and at least one group will block intervention (formally, (14) is satisfied for at least one group). Otherwise, choose a strong mandate.

8. Intervene if there is conflict and the global resource loss from intervention is less than the loss without intervention, that is, condition (1) is satisfied.

A fourth party

The model can be reinterpreted to allow a fourth party to intervene if M does not. We can capture this by viewing the no-intervention payoffs as the payoffs resulting when M does not intervene, even though another outside party may. For example, in deciding whether to intervene in the conflict in the Democratic Republic of the Congo in 1998-99, the UN should have wanted to consider the consequences of not intervening: other countries, such as Rwanda, Uganda and Zimbabwe, would. Where relevant, we can also interpret the payoffs with intervention by a fourth party as the result of cooperation or competition between multiple fourth parties. The fourth party, unlike M, may be able to change the wealth shares and take a share for herself. Nonetheless, the formal analysis is the same as
above; one should just think differently about the payoffs. A more nuanced model could allow both M and the fourth party to intervene, but I conjecture that the qualitative results would be a mix of the outcomes when M does and does not intervene. Returning to the Democratic Republic of the Congo example, other countries kept forces in the country at the same time as the UN, but it seems that each intervener obstructed the objectives of the other.

**Post-conflict aid**

Suppose that an outside party, such as an aid donor, is expected to commit funds $h$ to reconstruction after the conflict (Collier and Hoeffler 2004; $h$ could also be debt relief). This is effectively like a drop in $x$ or $y$, so by (4) or (5) it could promote conflict incentives _ex-ante_. In a slightly different framework, suppose that group 1 has increasing, convex effort cost function $c(e_1)$ and its wealth share is given by the function $s(e_1,e_2)$ when $s$ is increasing and concave in $e_1$. Then group 1 gets payoff $s(e_1,e_2)(\theta - d + h) - c(e_1)$. $d = x, y, \ldots$ and optimally sets $s'_1(e_1,e_2)(\theta - d + h) = c'(e_1)$, where primes and subscripts denote partial derivatives with respect to the variable in the subscript. The optimal $e_1$ is increasing in $h$. Likewise, if we allow both groups to choose effort then, in a symmetric Nash equilibrium, $\frac{1}{2}(\theta - d + h) = c'(e_1)$, and effort again increases in $h$. This can be shown to be true for many groups as well (Tornell and Velasco 1992). Thus, post-conflict aid can increase the incidence and severity of conflict by raising the gain to fighting (Keen 1998).
If donors can give a fixed amount of aid per group, independent of the group’s success in conflict, or even give aid which depends negatively on aggression and rewards groups which pursue their goals in non-violent ways, then aid is unlikely to encourage conflict or will even discourage it. However, avoiding that a group’s aggression raises the aid it receives after the conflict is complicated in practice because successful conflict participants can earn *de-facto* control of the country’s economy and wealth. Also, realistically, some donors may prefer to befriend the winner *ex-post*, causing time-inconsistency. While donors could also try to withhold aid altogether, the need for aid in the general population is particularly severe just after the end of conflict, so this policy too could be time-inconsistent. Also, again the donors may prefer to befriend the winner. However, since aggression is only one way to displease aid donors, insights from the broader aid conditionality literature may be available on this matter (Svensson 2000, Easterly 2002, Kletzer 2005, Azam and Laffont 2003, Torsvik 2005).

**Should interveners side with the winner?**

I have assumed that M’s intervention is unbiased to the extent it cannot manipulate the wealth shares of the parties to conflict. This may be either by choice (M only minimizes the total resource loss and does not pass judgment) or by political or military necessity. However, noting that impartial foreign interventions have often been unable to end conflicts (Regan 2002), Betts (1994) suggests that interveners should side with the winner to end the war quickly (see also Amegashie and Kutsoati 2007). However, beyond
the normative point that the winner may not be the most deserving party and the positive point that it may not be the best future ruler, this paper suggests the following objection: if M intervenes to help the winner, then the conflict may go from a proportional sharing outcome to a winner-takes-all outcome, which can further promote the incentive to fight.

Formally, suppose that with two groups M implements $s_i^1 = 1$ if $s_1^{ni} \geq 0.5$ and $s_i^1 = 0$ if $s_1^{ni} < 0.5$. This policy makes (4) more likely to hold if $s_1^{ni} \geq 0.5$ and otherwise, if $s_1^{ni} < 0.5$, the corresponding condition for group 2, $s_2^i(\theta - x) > s_2^{ni} \theta \geq s_2^i(\theta - y)$, is likely to hold, by $s_2^i = 1 - s_2^i$ (again, under realistic expectations). Thus, intervention is more likely to induce moral hazard because the gains to winning are so large. Siding with the winner makes no difference for incentives if intervention is not expected, and so it does not affect the incidence of the insufficient deterrence problem. Thus, although we can still pick a $\tilde{p}$ satisfying (12a) and (12b), and we may be able to do likewise in the case of many groups, unless siding with the winner significantly decreases $x$ (which could, however, raise the number of conflicts), a policy of siding with the winner seems unlikely to reduce global resource losses from conflict, and may do the opposite.

Like conflict aid when property rights to aid are a function of the success in conflict, siding with the winner can be a subsidy to aggression. On the other hand, however, if the intervener can commit to support the opponent of any aggressor, this will unambiguously deter conflict \textit{ex-ante}, although inevitable conflicts may become more protracted and costly. A policy of supporting the opponent, however, may lead a time-consistency
problem to reappear: *ex-post*, supporting the winner may lower the social loss or the intervener may find it in its own political interest to befriend the winner.

4. **Conclusion**

This paper has analyzed the incentive effects of different intervention policies and introduced the concept of insufficient deterrence. The insufficient deterrence problem is the opposite of the moral hazard problem in the sense that the absence, rather than the presence, of expected intervention is the cause of conflict. I have also sketched a procedure for the optimal design of intervention policy. Intervention policy should be set on a case-by-case basis and interveners must consider that, precisely when staying out can dissuade one party from starting a conflict, it can encourage another. It may be optimal for interveners to intervene with random probability. I have also argued that post-conflict aid and interventions that side with the winner, even if possible, can promote conflict by raising the return to fighting for at least one party. Thus, even if siding with the winner and post-conflict aid can lower the total losses for parties to unavoidable conflicts, such policies can raise the number of avoidable conflicts. To counter this rise, if time-consistency can be ensured, post-conflict aid should be independent or negatively related to a party’s aggression and interveners should side with the loser, not the winner.

Finally, I propose three caveats. First, under uncertainty about the parameters of conflict, the optimal design of intervention policy is complicated. For example, a commitment not to intervene can be costly if it turns out that intervention would have lowered the resource
loss or prevented conflict altogether. The optimal policy design under uncertainty
depends on the joint statistical distribution of the parameters of each conflict. Second, in
practice intervention policy may have to be designed not on a case-by-case basis but a
regional level. In this case, preventing a moral hazard problem in one country could lead
to an insufficient deterrence problem in a country close by. The optimal policy then again
depends on the distribution of conflict parameters (now for countries in the region).
Finally, while this entire paper has been concerned with preventing conflict, one should
keep in mind that, in the end, global welfare may be higher by allowing, for example,
conflicts which overthrow poor governments and quell ruthless rebels, rather than always
keep the peace.

1 As discussed in the special issue, there is debate on whether the term “moral hazard” is appropriate to
describe the problem. I think that it is indeed and will follow the precedent. As mentioned in Kuperman
(2005), moral hazard has long been used in economics to describe excessive risk taking by countries who
expect IMF rescue packages in case of financial crises (Lane and Phillips 2000 and references therein).
Though in some cases, the rebel groups discussed in the special issue may not be taking on excessive risk
themselves, they make the civilian population take on excessive risk because they benefit from the
insurance “paid out” to civilians when there is intervention.
2 For example, forceful US intervention stopped Serbian aggression in Bosnia, British intervention finally
ended the recent civil conflict in Sierra Leone, and failed interventions in Rwanda and Sudan may have
allowed the genocides occurring there.
3 Following Doyle and Sambanis (2006), I use peacekeeping to denote efforts to sustain an established
peace. Peacemaking is non-violent efforts to establish a peace. Peace enforcement is establishing a peace
by some degree of force, when at least one party to the conflict would otherwise obstruct the peace.
4 However, to resolve this moral hazard problem, condition (1) must not be satisfied, so \( y \) must not be too
large, unless \( M \) can commit not to intervene.
5 However, to resolve this insufficient deterrence problem, (1) must be satisfied, so \( x \) must not be too
large, unless \( M \) can commit to intervene.
References


*Ethnopolitics* (2005), 4(2), Special Issue: Moral Hazard.and Intervention.


**Appendix I: Coasean bribes**

By the argument in Coase (1960), widely known as the Coase theorem, in principle a defending group 2 could offer a transfer \( s_i (\theta - x) - s_i^m \theta \) or \( s_{1i} (\theta - y) - s_1^m \theta \) to an attacking group 1 to avoid the conflict. Such transfers to avoid war, however, often fail to occur in practice. One reason may be overconfidence on the part of either or both groups, so that, for example, groups 1’s estimates of \( s_i \) or \( s_{1i} \) may be upward biased. Alternatively, as discussed in Fearon (1995) groups may deliberately over-represent their estimates, there may be indivisible stakes, or, if we suppose that the wealth share
appropriated in war reflects military strength, and military strength reflects pre-war economic strength, then \( s_1' / s_1'^{wa} \) or \( s_1'' / s_1''^{wa} \) may be a constant ratio. In this case, a wealth transfer before the war improves both the no-war option and the war-option of group 1 proportionally and will not deter conflict. Group 1 then suffers from the problem that it cannot commit to stay peaceful in the future. Likewise, suppose that group 2’s assets are mainly human capital which cannot be alienated. Groups 1 and 2 may now prefer to negotiate a sequence of future transfers. Yet, in the future, group 2 may grow stronger and this would lower the share \( s_1' \) or \( s_1'' \) achievable for group 1 if group 2 reneges on the agreement. Now, a group 2 commitment problem leads to conflict (Wagner 2005). Finally, Garfinkel and Skapedas (2000) show how conflict can be preferable for both parties even under perfect information because the winner does not have to invest in military protection in the future. For some combination of these reasons, a Coasean contract may not be feasible. See Fearon (1995) for additional discussion.
Appendix II: Moral hazard, realistic expectations, and small $x$ and $y$ makes conflict inevitable for any number of groups

Moral hazard for small $x$ and $y$ implies that conflict is inevitable for any number of groups $N>1$, because there will either be insufficient deterrence for at least one other group $g$, or at least one group will find conflict irresistible regardless of intervention. To show this, I follow the steps leading to (10) and again impose realistic expectations. Then moral hazard for group 1 implies that, for the other $N-1$ groups,

$$\sum_{g \neq 1}^{N} s_{g}^{mi}(\theta - y) + y \geq \sum_{g \neq 1}^{N} s_{g}^{nu} \theta > \sum_{g \neq 1}^{N} s_{g}^{i}(\theta - x) + x,$$

(10’)

which means that, for $x$ and $y$ sufficiently small, either (i) the insufficient deterrence condition for at least one group $g$,

$$s_{g}^{mi}(\theta - y) > s_{g}^{nu} \theta \geq s_{g}^{i}(\theta - x),$$

(11’a)

must hold; or (ii) the parameters for at least one other group satisfy

$$\max\{s_{g}^{mi}(\theta - y), s_{g}^{i}(\theta - x)\} > s_{g}^{nu} \theta$$

(11’b)

which means that the group prefers conflict regardless of intervention. Thus, with $x$ and $y$ sufficiently small, conflict is inevitable.