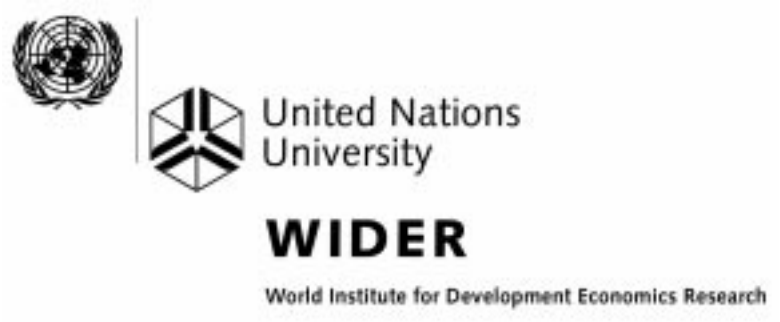
Addison, Tony; Murshed, S. Mansoob

Working Paper

Credibility and reputation in peacemaking

WIDER Discussion Papers // World Institute for Development Economics (UNU-WIDER), No. 2001/45

Provided in Cooperation with:
United Nations University (UNU), World Institute for Development Economics Research (WIDER)



Discussion Paper No. 2001/45

Credibility and Reputation in Peacemaking

Tony Addison and
S. Mansoob Murshed*

August 2001

Abstract

The paper analyses credibility and reputation in the context of peace negotiations. Where war provides economic gains to one side, peace is not incentive compatible, and peace agreements will necessarily degenerate, as they become time inconsistent. Levels of conflict are an increasing function of greed and rents, but decreasing in the direct costs of war. In a multiple period framework there is some uncertainty regarding the type of negotiator and for high values of the discount rate more conflict is chosen. Sanctions, aid and direct intervention, if effective, could eliminate conflict, as well as help in devising commitment technologies.

Keywords: peace agreements; credibility; reputation; conflict; aid.

JEL classification: O10; O55

* UNU/WIDER, Katajanokanlaituri 6b, Fin-00160, Helsinki, Finland; Addison@wider.unu.edu; Murshed@wider.unu.edu

This study has been prepared within the UNU/WIDER project on Why Some Countries Avoid Conflict While Others Fail? which is co-directed by Dr Tony Addison and Dr Mansoob Murshed.

Acknowledgements

We are grateful to Alex Cukierman, Ingrid Eide, Leif Helland, Halvor Mehlum, Kalle Moene and Prasanta Pattanaik for helpful comments and discussions. The usual disclaimer applies.

Introduction

Why do so many peace agreements fail? The model in this paper is applicable to the credibility of peace agreements in post-conflict situations in the immediate aftermath of civil war, where there is a danger of conflict breaking out again. It is not about the origins and causes of conflict *per se*, but about the difficulties in sustaining peace agreements. Our analysis is motivated by the empirical regularity with which peace agreements breakdown in the context of civil war. Also, several repeated attempts at peacemaking are usually necessary before lasting peace is established. Authors such as Walter (1999) have pointed out that the greatest impediment to the resolution of disputes is the inability of the belligerents to make credible commitments to peace. In turn the credibility of a peace offering or agreement hinges on the reputation of those making the offers. The upshot is that peace agreements are not sustainable when peace is not incentive compatible for one side entering into the agreement. Furthermore, when a bad reputation is inherited even genuine attempts to make peace will break down, as commitments are not credible.

Within the game theoretical literature, Azam (2000, 1995) models situations where the commitment to a peace agreement is important. This raises issues of credibility in connection with any peace agreement, which in turn depends upon reputation. There could also be the temptation to cheat on pre-announced commitments if the potential spoils of war are large enough. Besides the acquisition of power, a criminal motive to acquire valuable natural resource rents (diamonds in Angola and Sierra Leone) could result in reneging on a peace deal. For example, Collier and Hoeffler (1999) emphasise that greed, and not just grievance, is the main reason for many internal conflicts in Africa. They argue that quite often greed is concealed as grievance, which is an excuse to loot and expropriate. If that is the case then it might be in the interests of a warring party to make peace when external pressure is exerted, but renege at some convenient point in the future. It also means that genuinely peaceful overtures will be treated with suspicion.

The purpose of this paper is to construct analytical game theoretic models of peace agreements involving the credibility of pre-announced offers. We examine the incentive to renege on peaceful commitments, and how the reputation of a group extending the olive branch to its opposition might evolve over time. We also take into consideration the various temptations to renege on peace arising from the desire to acquire natural resource rents, “blood” diamonds for example. See Addison, Le Billon and Murshed (2000) on conflict motivated by natural resource rents, where poverty and injustice also play a part. Our purpose here is twofold: we show the difficulties of making peace credible even when intentions are good, and secondly we demonstrate the various temptations that might exist encouraging parties to renege on peaceful agreements. The peace agreement involves a commitment not to be belligerent that may not be tenable. Furthermore, we also speculate on how cheating on an announced peace deal could be minimised. As reputation is key to determining credibility, we analyse the formation of reputation in both single and multiple time periods.

The model that we apply has similarities to the monetary policy literature concerned with credibility and reputation in the context of inflation control. This is *because peace agreements* that seemingly end low intensity conflicts entail conceptually similar commitment problems to those faced by monetary authorities in achieving a zero or low inflation target. Reputation matters, and there is a temptation to renege on pre-announced commitments. Also, surprise war in the context of low intensity civil conflict brings about gains that are analytically akin to surprise inflation, and a multiple range of conflict equilibria are possible. But the policy implications and commitment technologies that we propose are considerably different from the monetary policy literature, as they can involve redistributive strategies in highly unequal post-conflict societies. We do not choose a repeated game framework leading to cooperation, as the risks associated with conflict re-emerging in most developing countries are associated with changes in leadership amongst the various parties.

We will analyse the reputation and credibility of only one side's announcements and offers. In our example it is the rebel side, but that role can be assigned to the government or another group without any loss of generality. Clearly, peace agreements involve at least two parties. But the problem of credible commitment usually involves serious difficulties for one side after a civil war. Where the dominant cause of the war is a grievance on the rebel side, such as historical economic discrimination (Rwanda, for example), then the credibility problem rests with the government offer, as in Azam (2000). When greed or natural resource rents are at stake, it is often the rebel side that is most recalcitrant and their commitment to peace treaties doubtful (Angola, Sierra Leone). Our paper is about the problem of commitment to peace, and the conceptual findings are equally applicable to any side whose self-interest is only dubiously compatible with peace. In other words, our model is about peace agreements in post-conflict societies where one side is weary of war, whereas the cessation of hostilities is not clearly incentive compatible for the other side. Examples of the former are ordinary people and their representatives; the latter group could consist of warlords as well as their henchmen. Ultimately war benefits only a few, and recent civil wars demonstrate that the majority of casualties are concentrated in the non-combatant civilian group (Azam and Hoeffler, 2000).

The rest of the paper is organised as follows. Section 1 is concerned with credibility and reputation in a single time period, analysing the temptations to cheat on agreements and how they might be tackled. Section 2 involves multi-period analysis, where reputation is not clear-cut with honest and dishonest types; we also look at the role of the discount rate in the decision making process. Finally section 3 concludes with a summary.

1 Credibility and reputation in a single period

The basic set up of the model involves two sides whom we refer to as government and rebels. As discussed above, one side is either tired of fighting or it has an interest in peace. This group is the government in our example. The other side, the rebels, may have something to gain from the resumption of fighting such as rents from resources. To reiterate, the roles played by the government and rebels in the games that follow can be reversed without altering the results. Both sides have entered into a peace deal. The government side derives no benefit from breaking this agreement. Consider the utility function of the rebel group (U^R):

$$U^R = -(1/2)c_1 w^2 + \theta c_2 (w - w^e) \quad (1)$$

where

$$\theta = B + \varepsilon$$

$$\varepsilon = \varepsilon_{t-1} + \eta \ldots \eta(0, \sigma^2)$$

The first term on the right-hand side of equation 1 is the pure cost of conflict in quadratic (squared) form, where w represents warfare or belligerent behaviour[1] and c_1 is the parameter measuring the direct cost of warfare. The negative sign before it is to indicate the cost or disutility from fighting. The quadratic form of the cost indicates that the costs of war rise more than proportionately as the level of w rises, implying that the low intensity conflict is 'less low'. The second term on the right hand side of (1) indicates the gains to the rebels from reneging on a peace agreement, or the benefit from a 'surprise' war, where the level of actual conflict (*w*) exceeds the level of conflict *expected* in advance (w^e). In other words, the spoils of war can only be wrested via the ruse of peaceful intentions. The parameter c_2 captures the magnitude of this effect, the higher is c_2 the greater is the gain from feigning to make peace first and looting later. It may also be viewed as a subjective measure of *greed*. In addition to this the greater the abundance of lootable resources, or rents to be extracted, the higher is the gain from surprise war. This is measured by the parameter θ which captures the rent (*B*) from disputed natural resources such as oil, diamonds and so on. It could include financial contributions from non-residents, and during the cold war period fungible military aid from superpowers. The rent or booty is subject to random shocks (ε) with a first-order auto-regressive process resulting in shocks persisting for some time. Random shocks could arise from terms of trade fluctuations or sudden increase in outside contributions to the war chest.

Equation 1 may be regarded as the utility function of the rebel leaders, and their followers whose participation and incentive compatibility constraints have to be satisfied to induce them to follow warlords. Furthermore, peace may actually reduce the utility of warlords, or a faction amongst them, who will then look for excuses to resume fighting. This often takes the form of an endless stream of grievances. Note also that the cost of fighting and gains from surprise war are additive separable.

As far as the gains from surprise war are concerned it is part of a process of income generation for the rebels (y^R) described as:

$$y^R = y^N + \theta(w - w^e) \quad (2)$$

Here the income of rebels is equal to some fixed or natural rate (y^N) plus an additional component arising from surprise warfare. The process described in (2) is similar to the Lucas aggregate supply relationship. The income associated with the natural rate is guaranteed and received with certainty. In contrast, the gains from the surprise element are based upon capturing the prize, and driving a wedge between actual and expected

---

1 The variable w may be measured in terms of chosen military expenditure and other costs of war broadly defined. We omit aspects of 'conflict technology' that are concerned with the trade-offs between peaceful and military production, as well as the probability of winning, as they are of marginal interest to the problems addressed in this paper. On these issues see, for example, Hirshleifer (1995).

levels of belligerence. It also means that if y^R is to be augmented above the natural rate, it has to come via surprise war. Hence only the second term in the right hand side of (2) is incorporated in (1)

The rebels maximise their utility in (1) subject to w, which leads to:

$$w = \theta c_2 / c_1 \tag{3}$$

This result can be interpreted in the following manner: the equilibrium choice of warfare is greater the higher is the element of pure avarice, c_2, the higher the availability of lootable resources and other sources of finance, θ, and the smaller the direct cost of fighting, c_1.

As far as the government (*G*) is concerned, a simple version of their utility function could look like:

$$\begin{aligned} U^G &= -(w - w^e)^2 \forall w \geq w^e \\ and &= (w - w^e)^2 \forall w < w^e \end{aligned} \tag{4}$$

The government's utility is declining in surprise warfare when actual war is greater than expected; however, it is the opposite when actual levels of war fall strictly below expectations. We will focus attention to cases applicable to the former, as the second line pertaining to instances were the actual levels of belligerency are below expectations (pleasant surprises) are largely irrelevant. The disutility from surprise war arises because the government has to engage in unforseen military expenditure that diverts income from other types of public expenditure or increases its borrowing/aid requirement. For example, Angola's government had to sharply raise military expenditure following UNITA's return to war in 1998.

The government is either tired of war or wants to enjoy its tenure in office in peace. In short, there are no gains from *surprise* attack or war for the government, and its utility declines in deviation of actual levels of war from expected levels (first line of (4)). We do not add any extra fixed costs of warfare in (4). Maximising (4) with respect to w^e yields:

$$w = w^e \tag{5}$$

The government reacts at the same time as the rebels. Substituting (3) into (1) for the rebel group, and (5) in to (4) for the government gives us:

$$\begin{aligned} U^R &= -(\theta c_2)^2 / 2c_1 \\ U^G &= 0 \end{aligned} \tag{6}$$

This is the outcome when the rebels have an incentive to renege on an announcement of complete peace, but it does not have a first mover advantage. Both announcements by the rebels and expectations formation by the government take place simultaneously. What if the rebels pursue a policy of no warfare with $w = 0$? Then:

$$U^R{}_P = 0$$
$$U^G{}_P = 0 \qquad (7)$$

This is the Pareto optimal outcome and superior to the result in (6). In the socially optimal state there is no war, and $y^R = y^N$. This is because the citizenry, who are the majority, suffer most from war. They may be goaded into supporting a war on the grounds of some grievance, but this typically cannot persist for long, as they begin to lose a lot from war.

Now if we assume that the rebels enjoy a first mover advantage and can announce complete peace and then engage in surprise warfare. In this case the actual and expected levels of warfare would diverge, $w = \delta c_2/c_1$ and $w^e = 0$ in equation 1. This involves cheating on a pre-announced commitment:

$$U^R{}_C = (\theta c_2)^2 / 2c_1 \qquad (8)$$

Note that the rebels' utility is greater in this case than under (6).

At this juncture we introduce reputation. Following Barro and Gordon (1983) the reputation of the rebels is all or nothing, and it hinges on its behaviour in the past. Consider the following rule. The opposition believes the announcement if the rebels' acted honestly in the previous period and kept its commitments. Otherwise it is not believed, and its actions are predicted to be that of a rogue group. This implies that there exists a future cost of cheating in the context of a low intensity conflict. The cost is equal to the loss of reputation and the inability to create surprises, but this cost (C) is in the future and is given by:

$$C = -((\theta c_2)^2 / 2c_1) \qquad (9)$$

Hence the penalty for cheating (which is the loss of reputation) appears to exactly equal the gain from cheating in (8). But the punishment comes in some future period. If the rebel group discounts this *future* loss, the cost of cheating is always less than the gain from reneging on a fixed commitment. Typically in conflict situations in many developing countries the future is heavily discounted. The upshot is that the socially optimal policy of zero warfare ($w = 0$) is *time inconsistent* or incentive incompatible, and thus will not be a possible outcome. The optimal policy of no conflict is infeasible, as it is not consistent with the incentives and expectations of the parties to the game. More particularly, the government knows it is in the interests of the rebels to renege on a pre-announced policy of total peace, and thus will not find any peace offering credible. Even groups with the best intentions in the world may be unable to negotiate a deal, as its peaceful overtures are just not credible.

Furthermore, there will be a range of possible conflict intensities that are feasible equilibrium outcomes. Thus multiple equilibria are possible. The results are depicted in Figure 1 in y^R and w space. The upward sloping linear aggregate supply curve has a slope exactly equal to θ, from equation 2, and is steeper the greater the availability of war booty or loot. The rebel's preferences are shown by the concave indifference curves with a slope = $\theta c_2/c_1$, obtained from (1). The greedier the group (the greater is δc_2), the steeper is the indifference curve. The rebels could announce zero conflict at point A. It could then cheat on its commitment and try to move to point B. The aggregate supply

curve schedule would shift leftwards because of the process of expectations formation. The vertical distance between B and C gives the range of multiple equilibria depending on the time horizon of the game and the discount rate used to obtain the present value of future reputation losses. The point B defines the lowest feasible rate of conflict. Following Barro and Gordon (1983), it is described as the *best enforceable outcome*, given the objectives of the rebels and the expectations of the government. To reiterate, a no war situation (point A) is simply not incentive compatible for the rebels, or credible to the government.

An increase in the available spoils of war, δ, shifts the aggregate supply function leftwards and makes the indifference curves steeper pointing to an expansion in the range of fighting. Note that this could arise due to random fluctuations on δ, and its lingering effects over time. Also situations where the future is heavily discounted (a low discount factor) are likely to raise the fighting threshold associated with the best enforceable outcome.

We now consider external policies to reduce conflict. Much of this implies manipulating the attitudes of the rebel leadership via sanctions, arms controls, trade restrictions and foreign aid. Consider a reformulated version of the rebel utility function where we embed external conflict prevention policy parameters, and an additional cost component associated with an implicit or explicit international agreement:

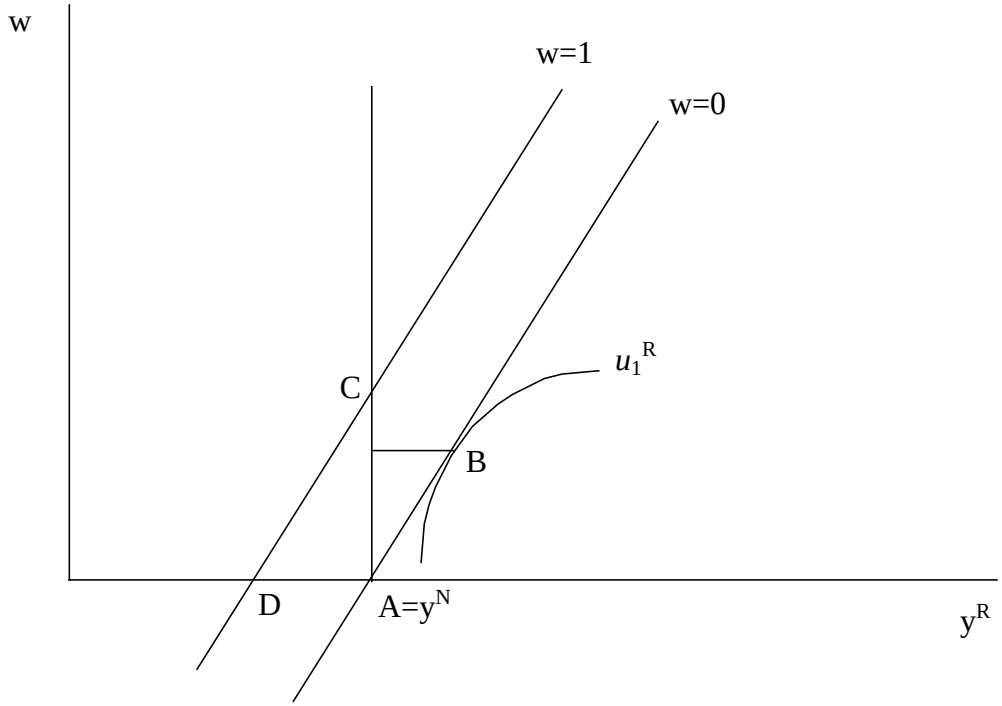
$$U^R = -(1/2)c_1(M)w^2 + \theta c_2(A,T)(w - w^e) - c_3(S)(w - w^e) \qquad (10)$$

Aid *(A)* may be utilised by a foreign power to reduce greedy attitudes, c_2. Similarly, trade sanctions *(T)* on items such as 'conflict' diamonds, money laundering and the activities of foreign entrepreneurs (in supplying arms and finance) might have the same effect (see Addison, Le Billon and Murshed, 2000). International controls on arms transfers *(M)* and/or sympathetic assistance from non-residents could be utilised to raise the direct cost of war, c_1. In either case, the indifference curve in Figure 1 will flatten out, and lower equilibrium ranges of fighting will emerge.

Another way of reducing conflict and belligerent behaviour is through an international agreement or understanding. Unfortunately, truly successful examples of these, such as NATO or the European Union, are exclusive to the developed world. Participation in these agreements implies a strategic pre-commitment to peace by delegation to a treaty or outside adjudicator. Although this method does not always eliminate conflict, it does raise the costs of war and reneging on peace deals. The last term in (10) represents such a *commitment technology* or *delegation*, and c_3 measures the costs of reneging on peace agreements as a function of sanctions *(S)* imposed by other signatories or parties to the agreement. Alternatively, the last term in (10) can be the force or delegated sanction imposed on warring parties by some outside power, such as an effective UN or another multinational peacekeeping force.

Figure 1

Credibility and reputation



Proposition 1: The presence of commitment technologies, sanctions, conditional aid and controls in international trade lowers the optimal level of belligerency amongst rebels.

Proof: Maximising (10) with respect to *w* yields:

$$w = (\theta c_2 - c_3)/c_1 \qquad (11)$$

This leads to a lower level of warfare when compared to (3). There is a direct effect of the external sanction or outside commitment technology, c_3. Then we have the indirect effects emanating from the manipulation of the behavioural parameters of the rebels. Chiefly, this involves aid and restrictions on the export/import of the rebels' international revenue/finance sources, which lowers c_2. Also, controls on arms transfers and financial assistance from sympathetic outsiders raise c_1. Note, however, that these policies involve costs to outside parties, typically borne by international donors.

A corollary of the above result is that the temptation to engage in conflict could even be completely eliminated ($w = 0$) if the "largesse", δ, is redistributed to become part of the natural rate of output in (2). Then the gains from capture and surprise war will vanish in rebel utility functions, as $\delta = 0$. It also means the policy innovations in (10) become

unnecessary, especially the last term ($c_3 = 0$). Such a redistribution of δ is more likely to be possible in cases where resource rents are more diffuse (agricultural wealth), rather than point-sourced like diamonds or oil (see Addison, Le Billon and Murshed, 2000). It also means that peace agreements, when reached, are more sustainable in countries with renewable and agricultural resources (Mozambique), than in countries with more easily capturable non-renewable natural resource wealth (diamonds and oil in Angola).

2 Reputation in a multi-period framework

Here the period of analysis involves more than a single time period. Also, the reputation associated with peace offerings may not be (0, 1), but something that evolves. Perceptions about reputation will be inherited from the past and updated using Bayes' rule. In contrast to the previous section there are two types of rebels. There will be honest *(H)* and dishonest *(D)* types, where the former is more dependable in keeping its commitments. A central finding will be that even dishonest groups, operating over a multi-period time horizon, may not break agreements at early stages of the game so as to leave their reputation intact for manipulation at later stages. This is similar to the Backus and Driffill (1985) result in the context of inflation control. But unlike in that model the discount rate, or the impatience to consume at present, could act against future reputational considerations. A generic objective function for both rebel types can take the following form:

$$U^{H,D} = -(1/2)c_1 w^2 + \theta c_2 (w - w^e) + \delta\left[-(1/2)c_1 w^2 + \theta c_2 (w - w^e)\right] \qquad (12)$$

Here we have extended the single period utility in (1) to two periods applying a discount factor, $\rightarrow$ to weight the future period. We do not include the policy extensions considered in (10) to avoid algebraic complexity. None of the qualitative results from above are altered except when specifically mentioned. Note that the discount factor, $\rightarrow = (1/1 + r)$, where r is an indicator of time preference. Thus the higher the discount factor the lower is the rate of discount. In other words, a higher value of $\rightarrow$ implies a greater concern for the future. Observe that we have reduced a multi-period game to two-period problem. This is because sub-game perfection allows solution by backward induction, and what matters are the penultimate and last periods, even when we extend the time horizon further.

It is instructive to examine decision making by the two types in the final period of analysis. A dishonest *(D)* type group in the second and final period of the game will simply choose the level of surprise warfare indicated by (3) in the previous section. This is because no discounting is involved in the last period, and the second term on the right hand side of (12) vanishes with $\delta = 0$.

The honest-type maximises utility (first-term on the right hand side of equation 12) with respect to two constraints (see Cukierman, 2000). The first is a dependability constraint; the honest type wants to appear to be true to its word:

$$w^H(2) = w^A(2) \qquad (13)$$

where the superscript *H* stands for the dependable or honest type, *H*'s announcements or offers are indicated by the superscript *A*, while the 2 in parenthesis indicates the second period. This constraint states that the actual outcome equals the announcement.

The other constraint concerns the government's beliefs about the type of rebel group. The government will assign a probability, γ, that the other side is the honest type and a probability $1 - \gamma$, that it is the dishonest type. Its expectation (indicated by the superscript e) of the level of conflict in period 2 will be a linear combination of the two strategies weighted by the corresponding probabilities:

$$w^e(2) = \gamma(2)w^A(2) + (1-\gamma(2))[\theta c_2 / c_1] \qquad (14)$$

Substituting (14) in (12), using (13), $\lambda = 0$, maximising with respect to w^H, yields:

$$w^H(2) = w^A(2) = (1-\gamma(2))[\theta c_2 / c_1] \qquad (15)$$

Note that the level of conflict picked by the *H*-type in (15) is lower than that chosen by the *D*-type in (3). Observe, however, that even the "better" type of group engenders conflict, as it is also non-altruistic. Knowing the group to be non-altruistic, the government will not regard over-optimistic levels of peace as a credible offer even from an honest type of group. Levels of w chosen in (15) vary proportionately with the poorness of equilibrium reputation, $(1 - \gamma(2))$. The result in (15) is akin to classic adverse selection problems in insurance markets, where the high risk type exerts a negative externality on the pooled contract offered to both the high and low risk categories.

Proposition 2: Uncertainty about the type of rebel leads to positive equilibrium levels of warfare even from the more honest or better type of rebel group.

Proof: In (15) if $\gamma(2) = 1$, $w^H(2) = w^A(2) = 0$.

If there was full separation of the two types of rebels, implying no uncertainty about the *H* or *D*-type, then $\gamma(2) = 1$ or 0. Otherwise in the presence of uncertainty, the government will use Bayes' law to update its prior beliefs about the rebel-type (see Cukierman (2000) for an exposition on how Bayes' law operates).[2] In this case, in period 2, we will have:

$$\gamma(2) = \frac{\gamma(1)}{\gamma(1) + (1-\gamma(1))\,p_1} \qquad (16)$$

This is the equilibrium value of the probability of the rebels being of the honest type in the second and final period of the game. It therefore captures reputational equilibrium. It also states that reputation in the second period is higher the greater it was in the first period and the lower is the probability of type *D* pretending to be *H* (p_1). Intuitively, this means a degree of path dependence or hysteresis. Thus some groups might wish to

---

2 Equation (16) is the posterior probability that the rebel is the H type, given that the H type has been played in period 1.

invest in a very good reputation in the initial period. This would lower their income drastically. Point D in Figure 1 would illustrate such an outcome.

Moving on to decision-making in the first or penultimate period, denoted by parenthesis 1, this is the period involving intertemporal choices over the two periods. This will mean discounting future utility levels, and choices about promoting and preserving a good reputation for use in future surprise warfare. Most importantly, the party making the offer will have to take into account the chance that its true type (either *H* or *D*) may be revealed in advance. This possibility implies that the other side knows for certain the type of group making the offer, an outcome known as "separation". By the same token, no separation *(NS)* means that the governments are left guessing about the rebel type that they are dealing with in the final period.

To develop the above point more fully it is worthwhile examining the utility levels of the two types of rebel groups in the final (or second) period under no separation and separation respectively. For the dishonest type under NS this is derived by substituting *w* from (3), w^e (2) = w^H (2) from (15) in the first two terms on the right hand side of (12):

$$U^{D}{}_{NS}(2) = \frac{\gamma(2)\theta^2 c^2{}_2}{c_1} \qquad (17)$$

With complete separation *(S)* the government knows that its opposition is of the dishonest type so that the second term on the right side of (12) vanishes, and there is no surprise war. All that is applicable is the first-term:

$$U^{D}{}_{S}(2) = -\frac{\theta^2 c^2{}_2}{2c_1} \qquad (18)$$

Note that for the dishonest type non-separation from (17) yields a greater utility than separation in (18). It will pay the *D* type to keep the government guessing about its type and therefore it will take care to conceal its true nature.

For the honest type rebel group utility in the second period with no separation involves inserting (15) into the first two-terms on the right hand side of (12), with $w = w^e$:

$$U^{H}{}_{NS}(2) = -\frac{[(1-\gamma(2))\theta c_2]^2}{2c_1} \qquad (19)$$

With full separation the above expression reduces to zero. Note that, unlike in the case of the *D*-type, the utility of the *H*-type is greater under separation (as the expression in (19) is negative). This is the converse of the case when we compare (17) with (18). Consequently the more honest type actually wants its type revealed to the opposition.

We can finally move to consider equilibrium in the first period. This will involve discounting the second period's utility. For the *D*-type utility will take the form:

$$U^D(1) = -(1/2)c_1 w^{D^2}(1) + \theta c_2 (w^D(1) - w^{D^e}(1))$$
$$+ \delta \left[\frac{-\theta^2 {c_2}^2}{2c_1} + \frac{\gamma(2)\theta^2 {c_2}^2 (w^H(1) - w^D(1))}{2c_1} \right]$$

The expression above is obtained from (12). The expression on the second line above corresponds to the third term in squared brackets in (12) and refers to the discounted value of playing the game in the second period. As such it contains two terms: the first term in line 2 of the equation above is equal to (18), and is the value of the game with separation; the last term refers to the value of the game with no separation, equation (17), multiplied by the probability of non-separation. This probability is: (w^H (1) – w^D (1))/2 for both *H* and *D* types (see, Cukierman, 2000 for details).

Differentiating the above with respect to $w^D(1)$, and rearranging terms yields the equilibrium level of conflict in period 1 for the D-type:

$$w^D(1) = \frac{\theta c_2}{c_1} - \frac{\delta\gamma(2)\theta^2 {c^2}_2}{2{c^2}_1} \qquad (20)$$

This demonstrates one of the most important results indicated at the beginning of this section. Even a dishonest type negotiator or rebel group will not engage in the full range of surprise conflict in the first period provided that there is a multiple period time horizon. This can be understood by comparing (20) with (3) above. The intuition is that a rational but selfish negotiator will want to keep his reputation intact for use in surprises at a later stage of the game. Another implication follows from (16) above, which states that the higher the initial reputation $\rightarrow(1)$, the greater is the second period reputation, $\rightarrow(2)$. When this result is combined with (20), it makes the more dishonest type less belligerent in the first period. Equally, if a bad reputation were inherited from the past, it would make this type of negotiator more reckless.

Turning now to the more honest type, the method involved in obtaining its utility function will be similar to that of the *D*-type above, using (12). In the second term of (12), however, the expression for the expectations of conflict in the first period *(w^e)* will take into account the government's uncertainty about the type of rebel it faces. This will turn out to be a linear combination of the conflict-weighted probabilities of being *H* or *D*. The honest rebel cares about dependability, dislikes misperceptions and wants its true type to be revealed. The expression in squared brackets contains the probability weighted utility of non-separation in the second period for the honest rebel given by (19). This is negative. The utility from full separation for the honest type in period 2 is zero. Thus:

$$U^H(1) = -(1/2)c_1 w^{H^2}(1) + \theta c_2 \left[(w^H(1) - \gamma(1) w^H(1) - (1 - \gamma(1)) w^D(1) \right]$$
$$+ \delta \left[\frac{-(1-\gamma(2))^2 \theta^2 {c_2}^2 (w^H(1) - w^D(1))}{4c_1} \right]$$

Once again differentiating the above with respect to $w^H(1)$ and rearranging terms gives us the first period equilibrium conflict level for the *H*-type:

$$w^H(1) = \frac{\theta c_2(1-\gamma(1))}{c_1} - \frac{\delta(1-\gamma(2))^2\theta^2 c^2{}_2}{4c^2{}_1} \qquad (21)$$

Thus, the honest rebel will pick a lower equilibrium level of conflict than the dishonest rebel. This can be demonstrated by comparing (21) with (20). Also intertemporal considerations will also make the honest type moderate initial period belligerency as well, which can be seen by comparing (21) with (15).

We now come to the central role of the discount factor. The higher is δ, the discount factor, the more important is the future, and less conflict is chosen by both types of rebel groups from (20) and (21). When the discount factor is low implying heavy discounting of the future even an honest rebel group will engage in high levels of conflict in the current period.

Proposition 3: When the future is heavily discounted, δ is very low, future reputational factors are ignored in favour of present consumption, producing greater levels of war.

Proof: In (20) and (21) as $\delta \rightarrow 0$, w (1) is higher.

In poor countries the future is heavily discounted (δ is very low), not so much for leaders, but their followers whose participation constraint in conflict is more readily met than warlords. Garfinkel and Skaperdas (2000) present a model where the discount rate could work in the opposite direction: caring about the future makes parties choose higher levels of conflict. This works because winning a conflict at present (destroying the opposition) increases future payoffs. This may be the case in gangland competition, some wars of secession, and in winner take all situations. A 'war of attrition' game takes place, where the presence of rivals in the game generates negative expected utility. But the discount rate is unlikely to work in the manner suggested by them in low intensity conflict, or where there is outside pressure from aid donors and regional powers to make peace. Concern for the future will lead to choices of low levels of conflict, as the opposition cannot be eliminated. In fact in poor developing countries conflict can arise from the relative lack of concern for the future. Present day survival, and current consumption dominate decision-making.

Finally, as far as the conflict prevention policy measures in (10) above are concerned, they will function in a similar manner, except that in the initial stages of a multi-period game the equilibrium levels of conflict chosen by rebels are lower. But it is the last period that really matters, and the results are therefore qualitatively unchanged.

3 Summary and conclusions

We have dealt with credibility and reputation in the context of peace negotiations. Although we refer to a rebel group the positions can be symmetrically reversed, with governments making the offers, and the rebels judging their credibility. Where war provides economic gains to one side peace is not incentive compatible, and peace agreements will necessarily degenerate. Socially optimal policies of no conflict may become time inconsistent with the expectation mechanisms in place. A range of equilibrium conflict levels may appear, implying multiple equilibria in a single period framework of reputation. The levels of conflict are an increasing function of pure greed

and lootable resources, but decreasing in the direct costs of war. Sanctions, aid trade restrictions and resource redistribution, if effective, might eliminate conflict.

In a multi-period framework reputation evolves, but initial reputation does matter, and there is some path dependence. We can also sub-divide the rebel group into relatively honest and dishonest types in the multi-period framework of analysis. The former is more interested in the revelation of its true type; the latter derives greater utility from concealing its intentions. Even a dishonest type negotiator will not engage in the full range of surprise conflict in the first period provided that there is a multiple period time horizon. A rational but selfish negotiator will want to keep his reputation intact for use in fooling his opponent at a later stage of the game. Of course, an honest type will pick a lower equilibrium level of conflict than his more dishonest counterpart. In multiple period analysis the less important the future, the greater is the chosen level of present conflict. Even a more honest type of negotiator will choose more current conflict for low (high) values of the discount factor (rate). This is indeed a serious problem in societies where poverty is endemic.

The first set of policy goals for conflict prevention concerns the economic cost-benefit calculus from fighting. Sanctions on the trade in arms, bank accounts and illicit natural resource based products will be useful. These include controls on the sale of 'blood' diamonds and restrictions on arms transfers to combatants. Collier (2000) also stresses the importance of stemming financial flows from Diasporas in the West to conflict societies in the developing world (an issue discussed further in Addison, Le Billon and Murshed 2000). Foreign aid can play a crucial part in reducing poverty and lowering the incentive to fight (as it did in the Mozambique's 1992 peace agreement). Foreign aid might also be used to make credible promises of redistributive reconstruction that close the gap in living standards between rebels and government supporters, thereby reducing the incentive to fight. There is also a role for the use of *force majeure*, military intervention by regional alliances and the UN. In order for this threat to work, military force must be credible, as is sadly not the case in many contemporary civil wars. In summary, conflict prevention by outside parties involves manipulating the utility function of combatants and their leaders so as to make fighting a less attractive prospect.

Thereafter, we have the problem of sustaining commitments to peace. Credible commitments are found in effective constitutional restraints and delegation. These domestic commitment technologies require institution building (Azam, 2000), something that is notoriously difficult to achieve. This is so because of the persistence of vested interests, and informal but powerful social networks.

References

Addison T, P. Le Billon and S.M. Murshed (2000). *'On the Economic Motivation for Conflict in Africa'*, Paper given at the Annual World Bank Conference on Development Economics, Paris, June 26-28th.

Addison T, P. Le Billon and S.M. Murshed (2000). *'Finance in Conflict and Reconstruction'*, Finance and Development Programme Working Paper 20, Manchester: Institute for Development Policy and Management.

Azam, J-P (2000). 'Civil Peace and Commitment to Ethno-Regional Redistribution in Africa', Paper given at the Annual World Bank Conference on Development Economics, Paris June 26-28th.

Azam, J-P (1995). 'How to Pay for Peace? A Theoretical Framework with Reference to African Countries', *Public Choice*, 83, 173-84.

Azam, J-P and A Hoeffler (2000). *'Looting and Conflict between Ethno-Regional Groups: Lessons for State Formation in Africa'*, Paper presented to a conference on The Economics of Political Violence, Centre of International Studies, Woodrow Wilson School of Public and International Affairs, Princeton University, and the Development Research Group of the World Bank, 18-19 March.

Backus, D and J Driffill (1985). 'Inflation and reputation', *American Economic Review*, 75, 530-8.

Barro, R and D Gordon (2000). 'Rules, discretion and reputation in a model of monetary policy', *Journal of Monetary Economics*, 12, 101-21.

Collier, P (2000). *'Policy for Post-Conflict Societies: Reducing the Risks of Renewed Conflict'*, Paper presented to a conference on The Economics of Political Violence, Centre of International Studies, Woodrow Wilson School of Public and International Affairs, Princeton University, and the Development Research Group of the World Bank, 18-19 March 2000.

Collier, P and A Hoeffler (1999). *'Justice-Seeking and Loot-Seeking in a Civil War'*, World Bank, unpublished manuscript.

Cukierman, A (2000). 'Establishing a reputation for dependability by means of inflation targets', *Economics of Governance*, 1(1), 53-76.

Garfinkel, M R and S Skaperdas (2000). 'Conflict without misperceptions or incomplete information: how the future matters', *Journal of Conflict Resolution*, 44 (6).

Hirshleifer, J (1995). 'Anarchy and its breakdown', *Journal of Political Economy*, 103(1), 26-52.

Walter, B (1999). *'Designing Transitions From Violent Civil War'*, IGCC Policy Paper # 31, University of California Institute on Global Conflict and Cooperation.

UNU World Institute for Development Economics Research (UNU/WIDER) was established by the United Nations University as its first research and training centre and started work in Helsinki, Finland in 1985. The purpose of the Institute is to undertake applied research and policy analysis on structural changes affecting the developing and transitional economies, to provide a forum for the advocacy of policies leading to robust, equitable and environmentally sustainable growth, and to promote capacity strengthening and training in the field of economic and social policy making. Its work is carried out by staff researchers and visiting scholars in Helsinki and through networks of collaborating scholars and institutions around the world.

UNU World Institute for Development Economics Research (UNU/WIDER)
Katajanokanlaituri 6 B, 00160 Helsinki, Finland

Camera-ready typescript prepared by Anna Kervinen at UNU/WIDER
Printed at UNU/WIDER, Helsinki

ISSN 1609-5774
ISBN 952-455-214-0 (printed publication)
ISBN 952-455-215-9 (internet publication)