Vicious Cycles: Endogenous Political Extremism and Political Violence

Rui J. de Figueiredo [*]   Barry R. Weingast [†]

VICIOUS CYCLES: ENDOGENOUS POLITICAL EXTREMISM AND POLITICAL VIOLENCE

Rui J. P. de Figueiredo, Jr.

University of California at Berkeley

Barry R. Weingast

Stanford University and Hoover Institution

Vicious Cycles: Endogenous Political Extremism and Political Violence

Rui J. P. de Figueiredo, Jr.
University of California at Berkeley

Barry R. Weingast
Stanford University and Hoover Institution

Version 1: September 1998 Current Version: October 2000

1. Introduction

Why do terrorists terrorize? Many explanations have been posited to answer this question. A common view, particularly in the popular press, is that terrorism and terrorists are *irrational*. A recent *New York Times* editorial attributed the violence experienced in Northern Ireland to 'hard-bitten loyalties and single minded devotion' which 'attracts psychopathic people' (*New York Times*, August 18, 1998: A23). Livingston (1978: 224-239) concurs that terrorists suffer from irrationalities and potentially even psychiatric illnesses. While such a view would make a rational account of terrorist activity seemingly impossible, in this paper, we argue that such an account is both possible and useful. Even in the event that terrorists suffer from such 'irrationalities,' a number of scholars have commented on the goal-directedness of those undertaking such terrorist campaigns.[1] And insofar as the pursuit of such goals is relevant, the kernel of a rational choice understanding of terrorism is possible.

In this paper, we take a first-cut at explaining terrorist (and anti-terrorist) behavior in societies with deep cleavages that might promote terrorism (for example, Northern Ireland, Israel, Sri Lanka, Indonesia, and India). Our purpose in this regard is three-fold: first to

---

[1] Gibbs (1989: 330), for example, in his discussion of how to conceptualize terrorism, includes goal-directedness as part of his *definition* of terrorism. "Terrorism is illegal violence or threatened violence," he

demonstrate that a rational account provides useful insights into the incidence and persistence of terrorism.[2] Second, our approach helps define conditions under which particular patterns of violence and terrorism—partly identified by competing explanations—prevail. And third, we make predictions about the type and character of interactions between terrorists and the majorities they seek to terrorize.

Beyond the 'irrational hatreds' argument that seems so pervasive, analysts have posited a number of possible explanations—implicitly rational—for terrorist activity. A variant of the 'irrational hatreds' story is that radicals simply gain a psychic benefit from hurting their opponents—from exacting *revenge* (cites). This argument has elements of instrumentality, albeit the approach is hard to justify from a strictly 'rationalist' perspective as it is so fundamentally preference-driven.[3]

Other analysts posit that terrorists seek to impose costs on the incumbent regime in order to force capitulation on the most important dimensions of cleavage. While often times rhetoric that incumbents are ripe to be overthrown belies the realistic probabilities that such a usurpation might occur, it is important to take this potential motivation seriously. One interpretation of such an argument is that terrorists *seek to impose costs* on the dominant group in order to extract concessions: in effect, they say, "I will stop hurting you, if you give me political concessions." Oots (1986) for example argues terrorists use this strategy as an exchange medium for "concessions." For such an argument to hold, however, there has to be sufficient credibility in terrorists' threats to continue terror in order for the threat to be effective and to stop terror is concessions occur.

A similar argument, albeit one we attempt to distinguish, is that terrorists terrorize in order to *provoke attacks upon themselves* (see for example Gibbs 1989; Lacquer 1987;

---

writes, "directed against human or nonhuman objects, provided that it:…(5) was perceived by the participants as contributing to the normative goal previously described (*supra*)…"

[2] In this sense, we build on the work of others who have also attempted to use rational explanations for political violence (see, e.g., Zinnes and Muncaster 198?; Sandler, Tschirhart and Cauley 1983; Enders and Sandler 1993; Lapan and Sandler 1988).

[3] As Fearon and Laitin (1996) point out, explaining ethnic political violence requires one to not only explain the cases in which ethnic conflict breaks out, but also the vast majority of cases in which it does not. In their estimation, violence is very rare in the universe of possible cases. Further, as others have added (de Figueiredo and Weingast 1999; Parikh and Cameron 2000), in order to explain such variation, one must appeal to explanatory factors that vary as well: fixed hatreds are insufficient. This is not to say, however, that hatreds are therefore necessarily excluded as an explanatory factor. Indeed, they might be a necessary (but not

Wilkinson 1986).[4] At first, this seems implausible: Why would a political group want to be attacked? The answer is political, rather than emotional. As Kropotkin wrote, "Terror is propaganda by deed" (in Rappaport 1984: 660). The logic of this claim is that highly public acts of terror often invite equally public or invasive acts of suppression. In so doing, the radical group encourages sympathy for their cause from more moderate, and usually larger, groups within their own group. This sympathy-generating aspect of terrorism is especially likely when either or both of two conditions hold: when the government cannot distinguish terrorists from the rest of the group and when their anti-terrorist policies impose costs on everyone, not just the terrorists. By encouraging these sympathies, radicals move portions of moderates from the radicals' own group closer to the radicals' policy perspective and thus encourage support for the radical political cause.

In this paper, we attempt to develop a theoretical model to help these views of why terrorists terrorize. In particular, our purpose is three-fold. First, the paper takes a first-cut at *positing conditions under which each of these explanations might hold.* In so doing, we hope to rigorously analyze a phenomenon which some have claimed has so far been impervious to theorizing.[5] As part of this exercise, the model helps to clarify the *interactions* which occur between these different theories, and in so doing, helps us generalize to *predictions* about what forms of terrorism (and suppression of terrorists) might obtain under different political-institutional contexts.

Second, in developing our model, we aim to explain some regularities that have been observed about regimes in which terrorism might occur. In particular, a rigorous theory of terrorism must not only explain cases in which terrorism arises as a persistent and dominant force, but also must explain the myriad of cases in which deep political cleavages *do not lead to terrorist activity.* Fearon and Laitin (1996), for example, document a number of cases in

---

sufficient) precondition—a proposition which is possible to subject to rigorous tests once interactions with other variables are deduced.

[4] Gibbs (1989: 339) summarizes this view: "Briefly, terrorists aim to provoke officials to such extreme repressive measures (e.g., censorship, preventive detention) that the government loses popular support and falls. The 'provocational' strategy is based on modulative social control, wherein the first party (terrorists in this case) uses the influence of the third party (the public at large in this case) on the second party (government officials in this case)." Later, in our model, we explicitly model the situation Gibbs describes—with three groups of varied power and therefore options—to explore the precise implications of his claims.

[5] "It is an illusion to suppose that social scientists have anything even approaching an adequate theory of terrorism," writes Gibbs (1989: 334). See also Lacqueur 1987: 165; Schmid and Jongman 1988.

which such deep cleavages do not lead to ethnic conflict, and similar claims can be made about terrorist activity (see also de Figueiredo and Weingast 1999).

Third, a similar challenge to theorists can be made on a temporal basis. States that have deep (ethnic, religious, economic) cleavages have the potential for terrorism, seem to exhibit a range of possible characteristics: sporadic episodes of violence, long periods of peace which suddenly erupt in extended periods of violence, or peaceful bargaining. To the extent that these are real types of 'violence regimes,' a theory must explain these patterns of violence and non-violence.

To examine these phenomena, we develop a model of radical political violence. The model has three players: a dominant ingroup (such as the Protestants in Northern Ireland or Israelis in the Middle East), and moderate and radical sub-groups of the outgroup (such as the Catholics or Palestinians). These players interact in two stages. In the first stage, radicals decide whether or not to terrorize the ingroup, which has the effect of imposing costs on that group. The ingroup in turn must decide whether or not to attempt to suppress the radicals, which has some chance of ending all future terrorism, but also causes the moderates of the outgroup to sympathize with the radicals more closely (in spatial terms, the moderates' preferences shift toward the radicals).[6] In the second stage, all three groups bargain over how much control and autonomy to cede to the outgroup. In the absence of a bargain, the outcome, we assume, is a civil war.

The game yields a number of interesting insights into the nature of radical politics and political violence. First, because we model the second stage as a complete information bargaining game, there is always some bargain that can avoid war. Further, the exact location of the bargain struck depends on the extremity of the moderates: the more extreme they are, the closer the final bargain is to the radicals preferred policy. This result sets the context for the first stage, for the radicals, looking to this future bargain, want the moderates to be as extreme as possible. The radicals' primary recourse is to terrorize and provoke attacks upon

---

[6] Note that some might argue that here the first stage should include the possibility of bargaining between the terrorists and the dominant group. Although this might be a reasonable assumption in certain cases, we exclude this possibility for two reasons. First, it allows us to include the moderates in the bargaining stage explicitly in the second stage (indeed, no temporal units are applied to the time period in which the game takes place). Second, and perhaps more importantly, as Corsi reports, for the kind of violence we are sicussing— bombing and assasinations as opposed to skyjackings and hostage taking—the perpetrators do not issue explicit *quid pro quo*-style demands in over ninety percent of cases (Corsi 1981: 60; Sandler, et. al. 1983).

their group by the ingroup so that the moderates will become more extreme. So why would the ingroup ever respond to such obvious provocation? The answer is that if the costs of future terrorism are high then they will be willing to risk alienating the moderates by attempting suppression. The model therefore sheds light on the nature of the interaction between radicals and the ingroup. One interesting point is that the incentives for the terrorists to terrorize and the ingroup to suppress terrorism diverge over time. As the radicals terrorize when the ingroup suppresses, they move the moderates closer and closer to themselves. This implies that the radicals' incentive to continue terrorism and thus to continue to bear the costs of suppression *decrease* over time. Once the moderates are sufficiently close to their position, the radicals will cease their violence. On the other hand, as the moderates move farther and farther away from the ingroup, the costs of suppression for the group (in terms of radicalization) also go down, meaning the incentives to suppress *increase* over time.

Second, this divergence means it is possible to divide regimes with the potential for violent activity into two classes: those in which at the outset, the ingroup will suppress and those in which it will not. Since the costs of suppression decrease over time, they will be their highest at their starting point. If the ingroup has an incentive to suppress at that time, it will always have an incentive to do so, at least until terrorism stops (either because the radicals have been suppressed completely, or they cease their violent activity). Alternatively, if a regime has no incentive to suppress, it will never do so, and terrorism will only be occasional and sporadic. The model therefore allows us to illuminate the conditions under which certain types of violent activity will take place. In particular, the model predicts conditions under which there will be sustained violence and suppression; and further, when such violence will stop. The model shows that persistent violence is most likely to be observed when moderates are truly moderate, in other words, not too close to either the dominant group or the radicals. The reason is that in this circumstance, both the ingroup and the radicals, have an incentive to engage in a tug-of-war for the moderate sympathies. When the moderates are close to the ingroup, the dominant group is not willing to risk their cozy relationship even if it means they must bear the costs of occasional terror. When the moderates are politically close to the radicals, the radicals have little incentive to terrorize, since there is scant room to radicalize the moderates further. Only when the moderates are between these two poles is it likely that ingroup's costs and the radical's benefits for

engaging in rounds of terror and suppression are most likely to *jointly* have the incentives to carry on cycles of violence .

Finally, the model highlights how the provocation motive interacts with other explanations to be a credible incentive for terrorist violence. In the first place, our results provide an explanation for why the ingroup would be goaded into radicalizing moderate outgroups: they will only do so if the threat of future violence is credible and the prospect of stemming that violence is sufficiently high. Further, the former implies that the radicals must have an incentive to terrorize *without* the provocation motive; otherwise the threat of future violence is not credible. Finally, the model highlights the fact that a provocation motive makes the incentives to terrorize much more powerful. The reason is that while costs imposed on the dominant group might provide utility to the radicals, this incentive is magnified dramatically when there is a political motivation as well: radicalizing moderates of their own group.

The paper proceeds as follows. In Section 2 we describe the model. Initially, we posit that the moderate outgroup radicalizes automatically upon ingroup suppression. We do this without positing the strategic microfoundations for such movement in order to establish general results. In Section 3, we describe the equilibrium to the model and explain the intuition behind it. In Section 4, we examine the dynamic characteristics of the equilibrium and attempt to characterize regimes in terms of these dynamics. In Section 5, we address the issue of why moderates might radicalize by developing an example of an incomplete information mechanism that provides similar intuitions to the reduced-form model presented initially. Finally, in Section 6, we offer some concluding remarks.

2. The Model

To analyze the nature of these phenomena, we employ the following model. The intention behind the model is to capture some of the most salient features of the strategic interaction characterizing political violence: an in and outgroup; political division between the outgroup; stages of violence and bargaining over policy; moderates who progressively radicalize if their group is attacked; and the potential for a vengeance motive by the radicals.

The model has two subgames. The first is the *terrorism game* (*TG*), the second, the *civil war game* (*CWG*). The two games have three players, M_1, M_2, and R_2, which represent the majority group, the moderate out-group, and the radical out-group, respectively. Each of these players are fully rational and therefore maximize their expected utility, with their utility functions consisting of two parts: one component based on the policy outcome, the second being the costs they incur during play. Thus, their utility functions are represented

$$U_i = -(p - x_i)^2 - C_i \qquad i = \{M_1, M_2, R_2\}$$

where *p* is the policy outcome, x_i is *i*'s ideal point and C_i is a general function which represents all of the costs incurred during play. The policy dimension can be interpreted as the collapsed dimension over which there is disagreement, such as rights or autonomy for the out-group, economic benefits, or other areas for divergence in policy ideals. Further, M_1's ideal point is assumed to be 0, R_2's ideal point is assumed to be 1, and M_2's ideal point is assumed to lie between these two, i.e. *x*∈ *(0,1)*.[7]

TG Structure. The structure of the *TG* is as follows (the extensive form is shown in Figure 1 and a summary of the action sets of the players is contained in Table 1). R_2 moves first and decides to either *terrorize* or *not terrorize* {*T,NT*}. If she chooses *NT*, the game moves to the bargaining game *CWG*. If she chooses *T*, there are three implications. First, it imposes a cost $C_t > 0$ on M_1. C_t is the disutility imposed on M_1 by the act of terrorism. Second, R_2 obtains a benefit ε>0. Finally, if R_2 plays *T*, M_1 then chooses to either *suppress* or *not suppress* {*S,NS*} the out-groups. Notice that playing *T* imposes costs only on others, and that R_2 does not incur any costs simply through the act (in fact, R_2 gets a benefit from the act itself). Of course, by allowing ε< 0, the model could accommodate such a disutility.

---

[7] Note that since M_2's ideal point is the only one that is fixed, we suppress the subscript hereafter and refer to her ideal point as *x*.

Figure 1. Extensive Form of Terrorism Game (TG)
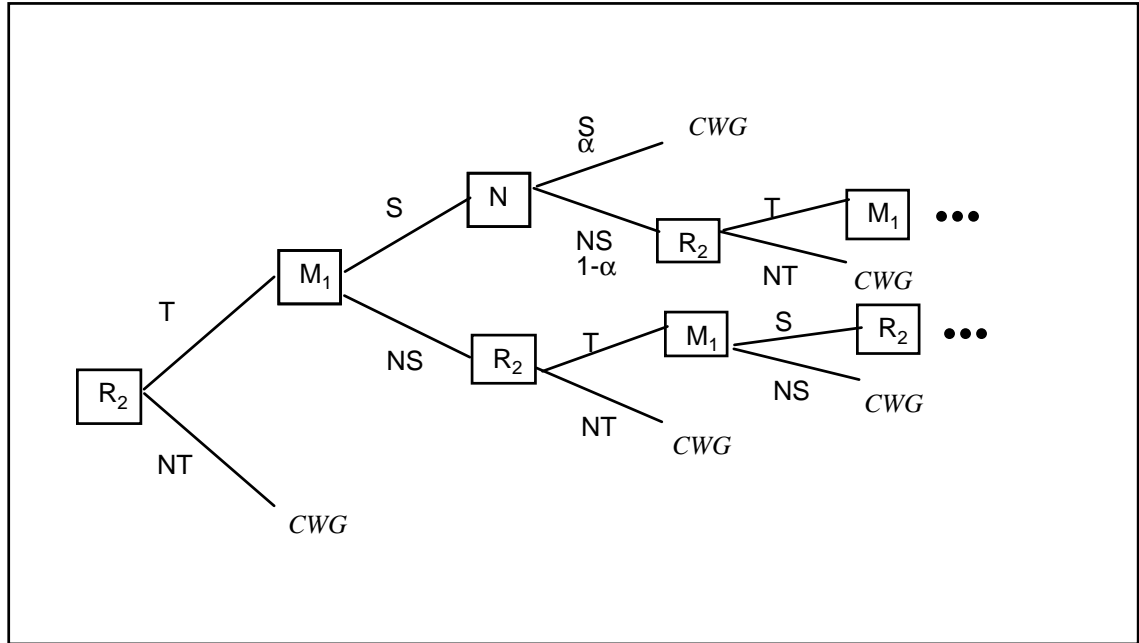
If M_1 chooses *NS*, then R_2 again chooses *T* or *NT*, with the same consequences as

before. If M_1 again chooses *NS* one more time, however, the *TG* ends, and a *CWG*

commences. The reason that R_2 is able to terrorize one more time is to reflect the fact that R_2

can terrorize during the bargaining period. Further, as the equilibrium analysis indicates, it is

possible to consider what happens if the game moves straight to a *CWG* if M_1 plays *NS*, an

interesting comparative static which will be explored more later.

If after R_2 plays *T*, M_1 plays *S*, again there are three consequences. First, M_2's ideal point

shifts closer to 1. We posit a specific functional form for this movement:

$$x' = (1 - (1 - x)\gamma)$$

where x is M_2's previous ideal point, x' her new one, and $\gamma \in (0,1)$.[8] Note that this means that

after a round of suppression, γ percent of the distance between M_2 and R_2 is closed. This

captures an important aspect of the nature of terrorist suppression. Suppression can turn a

---

[8] This model of preference movement might be considered a reduced form for an incomplete information game in which the moderates update about the ingroup's type based on their actions. With a bounded type-set, this would be sufficient to lead to an induced ideal point whose behavior was similar to x'.

moderate group more radical in one of two ways. First, policies to suppress the radicals might also harm moderate members of the out-group as well, as these two groups are often difficult to distinguish (such as in, say, Israel's 1996 attacks on Hammas targets which included civilians, British attacks on Catholics in Northern Ireland, or even during war, when civilian targets are mistaken for military strongholds). Even if effectively targeted, radicalization of moderates may occur from objections to acts of violence upon ethnically similar peoples, even if such violence or retribution is provoked.[9] Notably, by construction then, in this model one of the principal ways in which a radical group can influence an outcome is by *provoking* an attack upon *itself*. The second implication of M_1 playing S is that a cost $c_s > 0$ is imposed on R_2.

Finally, if S is played, the next move is made by a non-strategic player called *Nature* and denoted as player N. Upon M_1 playing S, then, N plays *successful* (S) with probability $\alpha \in (0.5, 1)$ and *not successful* (NS) with probability $1-\alpha$. The intuition behind this stage is that targeted retribution for terrorism meets with varying success depending on the case. Only rarely are radical groups wiped out, and in many cases, suppression of such a group is impossible for tactical and political reasons. Of course, the model accommodates such a case, simply by setting α to zero. If the suppression is unsuccessful, then, R_2 again gets to play T or NT, upon which M_1 plays either S or NS, and so on. This cycle of terrorism and suppression continues until either the suppression is successful, R_2 plays NT, or M_1 plays NS for two turns. In each stage that begins with R_2's choice of T or NT, the subgames are identical, except for the important feature that M_2's ideal point has possibly moved if there has been a previous suppression of the out-group.

A few comments are necessary concerning the structure of the *TG*. First, M_2 does not have a role in this stage. In fact, M_2 is interpreted to have a passive role, awaiting some resolution of the cycle of violence in order to take an active role. Second, the *CWG* commences, and perhaps more importantly, the cycle of terrorism and suppression stops, if one of the following occurs: M_1 successfully suppresses R_2; R_2 chooses to not terrorize; or M_1 does not suppress twice successively after terrorist acts.

---

[9] We distinguish between these two by calling the first *targeting error* and the second *triggering sympathies*.

Figure 2. Extensive Form of Civil War Game (CWG)
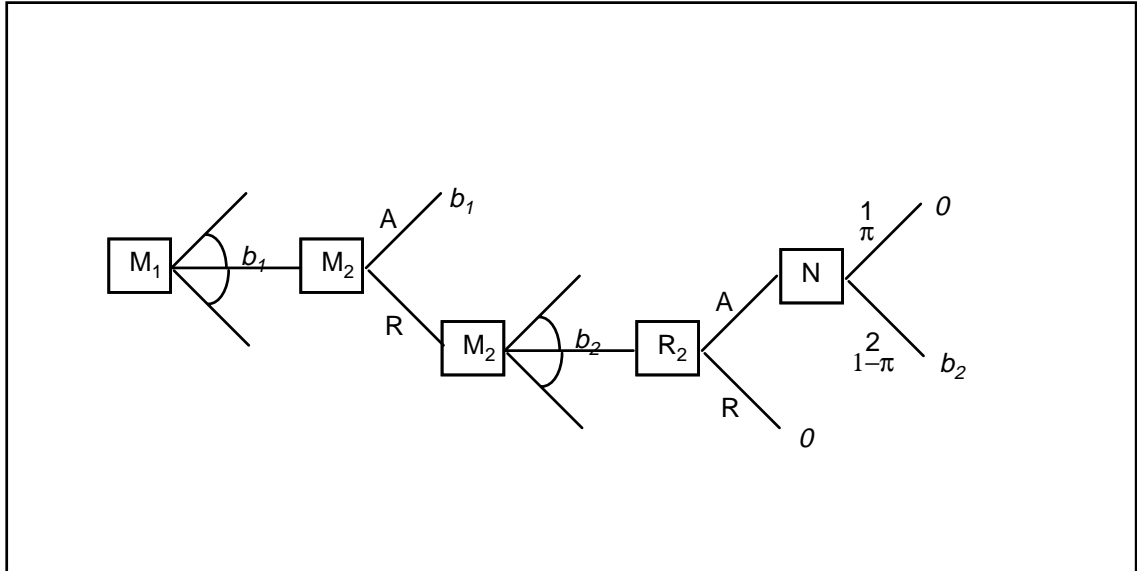
CWG Structure. The *CWG* is a *bargaining game* (the extensive form is shown in Figure 2). M_1 moves first and makes an offer $b_1 \in (0,1)$ to M_2. M_2 then chooses either to *accept* or *reject* the offer $\{A,R\}$. If M_2 accepts the offer, then $p = b_1$. Thus, the coalition of M_1 and M_2 is enough to reach a stable peace. However, if M_2 plays R, then M_2 makes an offer $b_2 \in (0,1)$ to R_2. R_2 then also must choose to either *accept* or *reject* the offer $\{A,R\}$. If R_2 plays R, then $p=0$; M_1 gets to impose its own ideal point. If R_2 accepts the proposal b_2, however, it is not certain that the coalition of M_2 and R_2 will triumph. Instead, there is political violence, with the out-group conducting a civil war against the in-group M_1. In this case, N chooses a winner, where M_1 wins with probability $\pi \in (0.5,1)$, and the M_2-R_2 coalition wins with probability 1-π. If M_1 wins, then $p=0$, and if the coalition of M_2 and R_2 wins, then $p = b_2$. One point of note about the *CWG* is that it might be considered that there should be costs imposed on all of the participants for fighting a civil war. This is certainly a modeling option. At this point, however, we forgo this option since the results are fairly clear: it will add an additional, albeit important, parameter to the model which will work in a linear fashion; thus if the costs of war are prohibitive, it will limit one or the other players' actions. However, since we are

interested in cases in which civil war is a viable strategy for one of the players, we limit attention to policy outcomes in this case.

Table 1. Summary of Action Sets in *TG* and *CWG*

Player	Action Sets	Description of Choice
R_2	$\{T,NT\}$	terrorize or not terrorize
	$\{A,R\}$	accept or reject b_2
M_1	$\{S,NS\}$	suppress or not suppress
	b_1	offer to M_2
M_2	$\{A,R\}$	accept or reject b_1
	b_2	offer to R_2
N	$\{S,NS\}$	suppression successful (prob α) or suppression unsuccessful (prob $1-\alpha$)
	$\{1,2\}$	winner of civil war is in-group (prob π) or out-group (prob $1-\pi$)

3. Equilibrium

What type of behavior results in a situation where the above assumptions prevail? To analyze this question we employ a solution concept known as subgame perfection. In particular, every actor is assumed to maximize his expected utility at every node, given the other players' strategy, and the equilibria derived for the overall game must also be equilibria in every proper subgame. An *equilibrium* in this game is a set of actions specified for every player which includes those which occur off the equilibrium path.

Using this solution method, we obtain a *unique* subgame perfect equilibrium for the combined *TG-CWG meta-game*. This equilibrium is stated in Propositions 1a and 1b and the proofs are contained in the appendix.

PROPOSITION 1A. The unique *subgame perfect equilibrium (SPE)* for the *TG* is as follows:

a. M_1 plays S iff $c_t > c_t^*$ where

$$c_t^* \equiv \frac{\alpha^2 (1-\sqrt{\pi})}{2\alpha - 1 - (1-\alpha)^{t^*+1}} \left[\frac{1-(1-\alpha)^{t^*} - x^2}{\alpha} - 2\frac{\gamma - (1-\alpha)^{t^*}\gamma^{t^*+1}}{1-\gamma(1-\alpha)} + (1-x)^2 \frac{\gamma^2 - (1-\alpha)^{t^*}\gamma^{2(t^*+1)}}{1-(1-\alpha)\gamma^2} \right]$$

where x is the induced value and t^* is defined by the maximum remaining rounds of terror (see appendix).

b. R_2 plays T iff

    (i) $c_s < c_s^*$ or

    (ii) $c_s > c_s^*$ and $c_t < c_t^*$ where

$$c_S^* = [1 - x(1-\sqrt{\pi})]^2 - [1 - (1-(1-x)\gamma)(1-\sqrt{\pi})]^2 + \varepsilon$$

PROPOSITION 1B. The unique *subgame perfect equilibrium (SPE)* for the *CWG* is as follows:

a. M_1 offers
$$b_1 = x(1-\sqrt{\pi})$$

b. M_2 plays A iff
$$(1-\pi)x^2 - 2b_1 x + b_1^2 < 0$$

c. M_2 offers
$$b_2 = x$$

d. R_2 plays A iff
$$b_2 > 0$$

How can one interpret the equilibrium defined in Propositions 1a and 1b? The first thing to note is that the SPE of the *CWG* generates a result in which the optimal offer b_1 will always be accepted by M_2. This has a number of implications. First, it means that in this model, *civil war never occurs*— M_1 will offer b_1^* and M_2 will accept it. This result generates a concern, therefore, for it is certainly the case that civil wars can occur in reality. In our model, because the positions of all of the players' ideal points are always known with certainty, it means that an optimal, war-avoiding bargain can be struck. If we include incomplete information about M_2's ideal point, as has been done more generally concerning the outbreak of war elsewhere (see, e.g. Powell 1989) it is the case that a civil war can occur . The fact that the outcome of the *CWG* is deterministic also makes analysis of the *TG* more straightforward: all of the parties know what the outcome of the bargaining stage will be, so it is possible to substitute $b_1^* = x(1 - \sqrt{\pi})$ at those nodes in the *TG*.

A second point illuminated by Proposition 1b is the nature of the bargaining power between the groups. Notably, since in this stage of the game, R_2 only has recourse to $p=0$ if it rejects M_2's offer, R_2 is relatively powerless to reject an offer from M_2. Therefore, it will accept any offer in which it could obtain an expected outcome greater than zero. Thus, M_2 will always offer her ideal point since it is strictly greater than zero; M_2 has no need to offer anything other than its ideal point.

A third point concerning Proposition 1b is that the bargain struck between M_1 and M_2 in equilibrium is affected by both the location of M_2's ideal point and M_1's chance of winning a civil war. These two factors operate in opposite directions. Since $\dfrac{\partial b_1^*}{\partial \pi} < 0$, this means that the less chance that M_2 has of winning a civil war, the more willing it is to accept a bargain which is closer to zero and, therefore, farther from its own ideal point x. This accords with intuition, since it means that the more powerful M_1 one is, the better a bargain it is able to extract. Similarly, since $\dfrac{\partial b_1^*}{\partial x} > 0$, as M_2 becomes more extreme in her position, the bargain that it obtains is more extreme. This is because under any conditions, M_1 will avoid a war; but to do so, it must offer M_2 a bargain which is relatively close to M_2's ideal point. As M_2 moves towards 1, then, so does b_1^*.

The equilibrium of the *TG* is not so uniform. Take R_2's decision to terrorize or not. If M_1 is not suppressing, then it will always be in the interest of R_2 to terrorize. The rationale for this is that terrorizing only gives positive utility to R_2, at no cost. Thus, if there is no chance of incurring the costs of suppression, it will play *T*. If M_1 is suppressing, however, it makes the trade-off between bearing the cost of suppression c_s, and moving M_2's ideal point closer to 1. Since the bargain struck in the *CWG* is a deterministic function of *x*, the closer that *x* gets to 1, the better the deal is for R_2. The cutpoint c_s*, then represents this tradeoff—if M_2's ideal point, and thus the eventual peaceful bargain, will move enough from another round of suppression, it will be greater than the costs of suppression to R_2, and R_2 will terrorize. If this is not the case, then R_2 will cease playing *T*. Further, since $\dfrac{\partial c_s^*}{\partial x} < 0$ as *x* increases, c_s*

decreases. This means that it becomes more and more likely that R_2 will not terrorize, the closer and closer that M_2's ideal point gets to 1. From this point it follows that the more that terrorism occurs, the less likely it is to continue. When the degree of movement of M_2's ideal point (which is a decreasing function of *x*) is small enough, she will no longer want to allow M_1 to suppress her by terrorizing. Note that it is possible if c_s is small enough, that terrorism will never stop, an intuitive result, since as c_s goes to zero, there is no disincentive to terrorize at all.[10]

M_1's decision to suppress or not is similar to R_2's suppression decision. As the proposition states, M_1 will only suppress if the cost of future terrorism is high. M_1's decision to suppress has two costs: first, it moves M_2 farther from M_1's ideal point, and thus affects the eventual bargain that will be struck; second, it invites future terrorism from R_2, which also could be costly. So why would M_1 ever play *S*? The answer is that M_1 trades a lottery over being terrorized and never being terrorized again, conceding some movement in her eventual bargain, against terrorism with certainty for one turn. Indeed, as $\alpha \to 0$, $c_t^* \to 0$, meaning that as the promise of successfully suppressing terrorism declines, it becomes an increasingly

---

[10] Note also that c_s has a number of other properties which we would expect—namely $\dfrac{\partial c_s^*}{\partial \pi} < 0$ and

unattractive option, and M_1 will not suppress ever. If α is high enough, however, M_1 will suppress. Further, under fairly general conditions, M_1's cut point c_t^* is decreasing in x, so as x increases, M_1 becomes more and more willing to suppress the outgroup.[11] The rationale follows directly from the ingroup's static incentives: namely as x moves closer and closer to 1, that portion of the suppression cost—namely the fact that it will alienate the moderates— becomes less constraining, since less of the distance between the moderates and radicals is left to close.

Another way to see the logic behind the equilibrium is to analyze each condition in Proposition 1a. The first set of conditions (conditions (a)) state that M_1 will suppress when the cost of terrorism is high (condition (i)) *and* the cost of suppression is not high enough to deter R_2 from terrorizing (condition (ii)). In other words, they will forgo suppression if radicalizing moderates is more costly than the expected benefit of (successfully) suppressing all future terrorism. The second set of conditions gives a similar calculus for R_2. R_2 will not terrorize once the cost of suppression, which they know will happen if the costs they impose are high enough, is higher than the benefit they get from suppression (i.e. moving M_2 closer to 1).

What does this mean, then, when we put all of the players' behavior together? The result is that for a range of starting values of x and other parameter values, there will be a *cycle of terrorism and suppression* until one of two things happen. First, M_1 might successfully suppress R_2, leading to a bargain being struck between M_1 and M_2. Second, the costs of suppression might be higher than the benefit from moving M_2's ideal point, so R_2 might not terrorize anymore (this behavior is illustrated in Figure 5). Otherwise, M_1 and R_2 continue to attack each other in their available means and the violence continues.

4. Dynamic Implications and Predictions

$$\frac{\partial c_s^*}{\partial \gamma} > 0.$$

[11] Specifically, as long as $x > \dfrac{\gamma}{1+\gamma}$, c_t^* is decreasing in x. This condition means that is γ is "not too large" relative to x.

In the previous section, we explained the logic behind equilibrium behavior. In this section, we take a deeper look at the dynamics of the equilibrium as a function of parameter values.

Figure 3 summarizes the logic of the players' one-turn calculus, given the other players' strategies. As Figure 3 illustrates, the most binding constraint is whether or not the ingroup will suppress or not. Only upon suppression, does terrorism continue for longer than one period. So when will there be an incentive for the ingroup to suppress terrorists upon acts of terrorism? It depends on two parameters: the cost of terrorism and the probability of successfully suppressing terrorists. If the costs of future terrorism are low, the ingroup would prefer not to suppress, since another round of terror is worth the benefit of more supportive moderates. Similarly, suppression is also not an attractive strategy for the ingroup when its likelihood of success is relatively low. In the limit, when the probability terrorists will be successfully eliminated is zero, there is no reason for an ingroup to pursue that strategy. Again, that leads to a system in which the there might be occasional terrorism, but in which cycles of violence cannot be sustained.

**Figure 3. Outcomes as a function of costs and α**
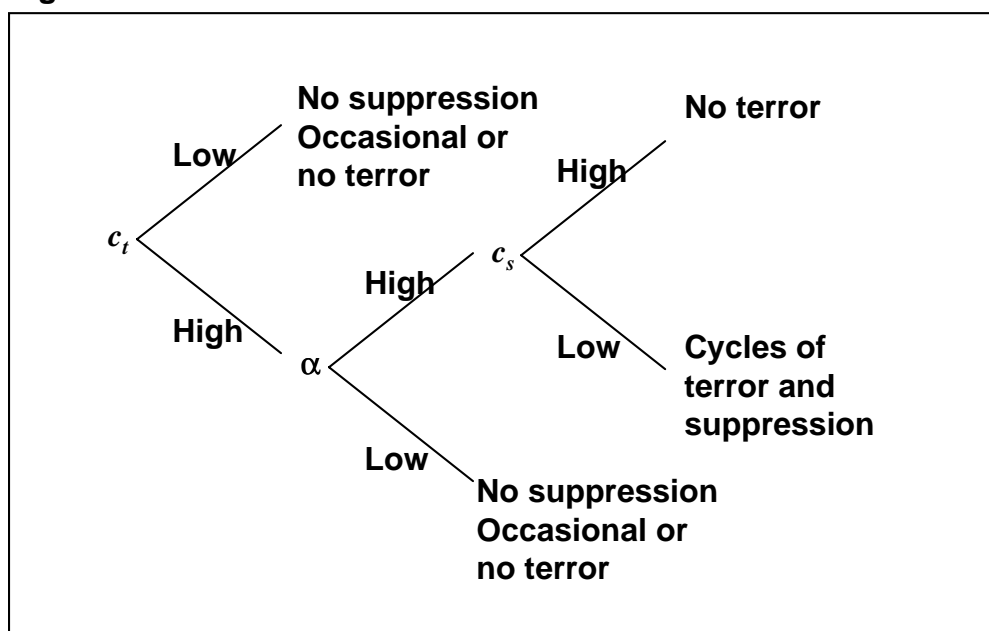
If suppression is a viable option, the strategy of the terrorists becomes the key determinant in the pattern of violence and suppression. Here, as noted earlier, the decision depends on whether or not the costs of suppression are sufficiently high, in comparison to the benefits the terrorists get in moving the moderates closer to their position. If the costs of suppression are high, no terrorism will result, since despite the political motive terrorists will not be willing to bear the pain of being suppressed. If these costs are low, however, then radicals will have an incentive to commit acts of violence, since the political benefit outweighs the short-term loss suffered by suppression. This then implies that in situations in which persistent terrorism results, the costs of terrorism and suppression must be low and the probability of suppression must be relatively high.

Moving to a more dynamic analysis, one of the questions we initially posed was when does a cycle of terrorism start and when does it stop? Consider first the onset of terrorism. As we mentioned previously, if the costs of terrorism are low the ingroup will not suppress the outgroup and the political incentives for terror decrease. At the same time, we noted that as x gets larger, the cutpoint also goes down, since the costs in terms of moving the moderates is relatively lower than previously, with each successive suppression.[12] Putting these two facts together means that the ingroup is always going to have the same response to terrorism for the entire play of the game. Figure 4 illustrates this point. If the ingroup has an incentive at the beginning to suppress terrorism, in other words if $c_{t0}^* < c_t$, where the *0* subscript means the initial value, then because c_t^* is increasing in x and x is weakly increasing in time, c_t will always be greater than c_t^* irrespective of the rest of the play of the game. So for the entire play, M_t's response to terrorism will be suppression. Alternatively, if $c_{t0}^* > c_t$, then in the first turn, when the radicals terrorize, the ingroup will not suppress them. This means that the ideal policy will not move, and thus the costs of terrorism will remain too high to provide an incentive to respond to future terrorisms. The implication is that it is possible to characterize regime types by their starting cut-points: those regimes which will always respond to

---

[12] Recall that for $\dfrac{\partial c_t^*}{\partial x} < 0$, $x > \dfrac{\gamma}{1+\gamma}$. If this condition does not hold, however, it does not imply that the dynamics discussed here will not result. For suppression to be a dominant strtaegy under any terrorism when the condition does not hold, it simply means that c_t must be greater than the maximum c_t^* given x. Otherwise, cycles of violence could be stopped by the ingroup.

terrorism, and those that will never respond to terrorism. Indeed, in this sense, after the first round, the stopping point for continued terrorism is fully determined by the incentives of the radicals.

**Figure 4. Illustrative Equilibrium Behavior in the *TG:* Incentives of *M_1***
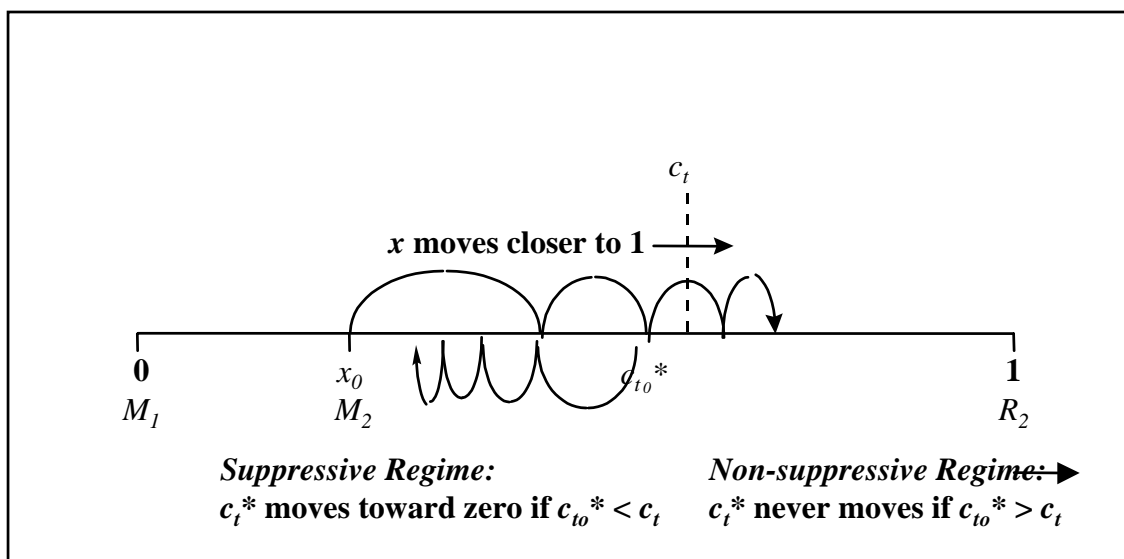
So how do the radicals incentives change as the game progresses? As we noted before, in a non-suppressive regime, their incentives to terrorize are restricted solely to the benefit they get from committing the act. In a suppressive regime, this incentive is augmented by the political incentive to try to move the moderates. Assuming that this is sufficient to justify the costs of suppression, when would they stop? As shown in Figure 5, when the radicals terrorize a suppressive regime, they will invite suppression, and will succeed in moving the moderates closer to the radical position. What happens if suppression is unsuccessful? In the next period, the incentive to terrorize goes down, for the same reason the costs of suppression for the ingroup goes down: the marginal gain from movement of the moderates' position becomes smaller as the moderates approach the radicals' position. This then provides an intuition about when the cycle of violence, having started, will stop: when the costs of suppression ultimately are higher than the radicals' cutpoint, terrorism will cease and the interaction moves to bargaining.
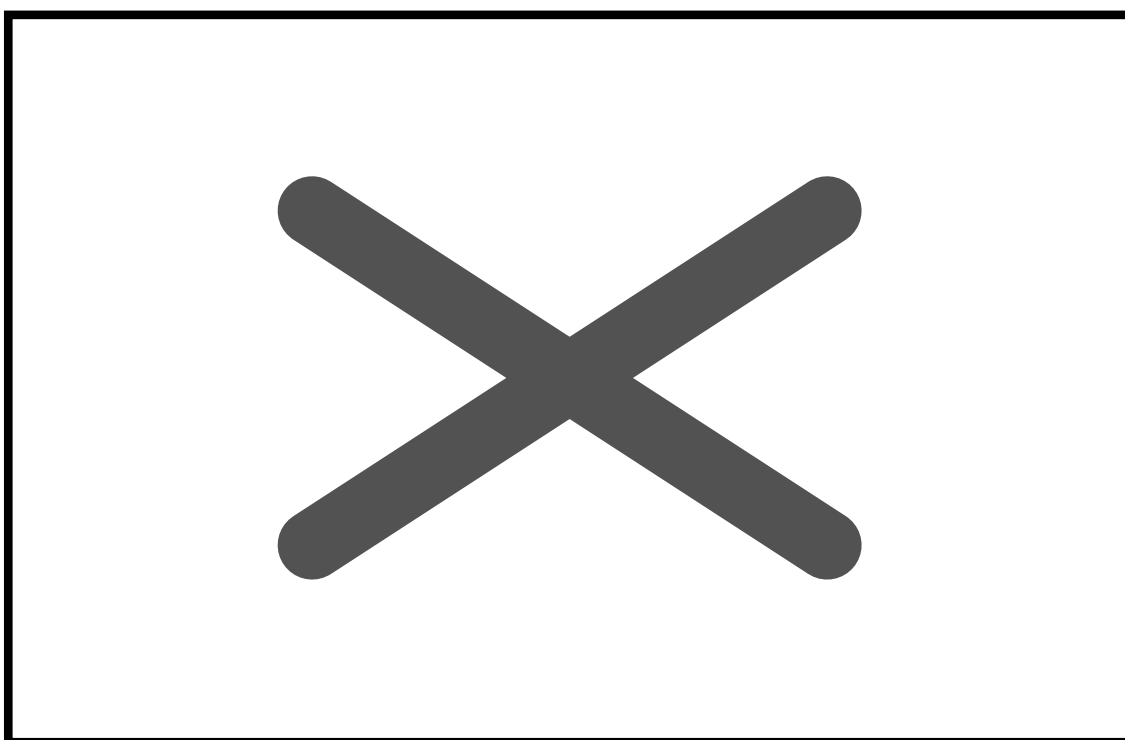
This suggests a final interesting way to characterize systems in which political violence might obtain. In particular, a key determinant of what type of dynamics a system will exhibit is the *starting position of the moderates*. Consider the incentives of each of the players as moderates' initial position moves from zero to one. In Figure 6, we characterize the system outcomes, having fixed the exogenous costs of suppression and terrorism on their respective targets, in terms of the probability of successful suppression and the initial position of the moderates. As Figure 6 illustrates, when x starts out in a moderate range, controlling for other exogenous parameters, sustained, cyclic violence is likely to obtain. The logic is as follows. When x_o is extremely low, the costs to the ingroup of suppressing are high (in other words c_{t0} is very low), meaning they will be unlikely to be suppressive. This limits the incentives (and ability) of the radicals to terrorize. Alternatively, when the initial position of the moderates is extreme, the ingroup might indeed be willing to suppress, but the radicals' incentive to invite such suppression is low; so for different reasons, a non-terrorism result obtains: in this case, it is the radicals who have little incentive to commit acts of political violence. When x_o is in its middle ranges, there is the potential for a convergence of

incentives to suppress and to terrorize given suppression. It is in these cases that one is most likely to observe persistent cycles of terrorism and suppression. Thus, *only when moderates are truly moderate, are cycles of political violence likely*.
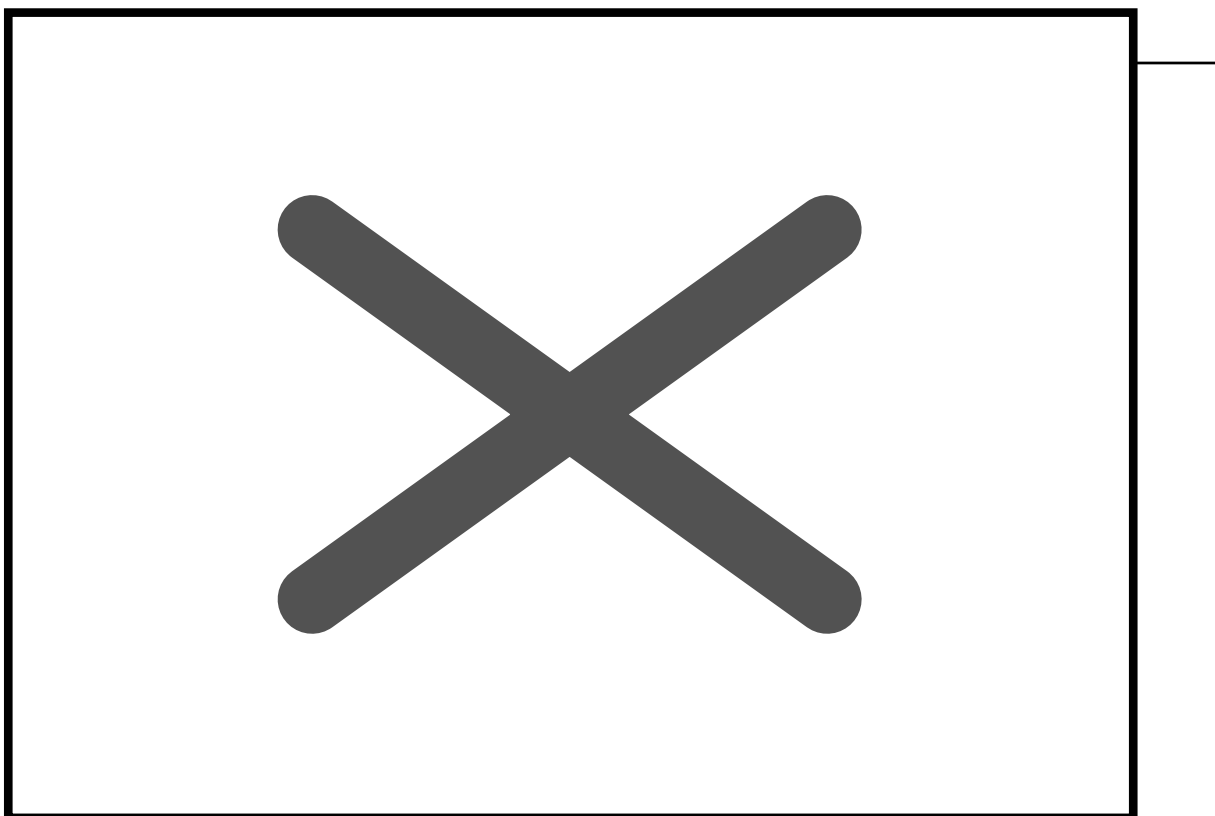
5. The Microfoundations for Moderate Radicalization:
An Incomplete Information Extension

Previously, we made two crucial assumptions about the moderate out-group's behavior in the *TG*: that the induced ideal point of the moderate out-group moves *automatically* and at a *decreasing rate* towards the radicals when the in-group suppresses. This reduced-form assumption can accommodate a number of possible more general incomplete information models, but is unsatisfying in that it shifts attention away from the strategic nature of the moderate's role in the early stage, and because it confounds induced preferences over outcomes given information and beliefs (which may shift) and underlying preferences or utility functions (which should not shift).

In fact, a number of mechanisms might provide the microfoundations for the radicalization behavior of the moderates. In this section, we provide an example of an informational structure which generates the dynamics observed previously. This model is more general in the sense that all actors behave fully strategically. Notably, in another sense, it is more restrictive since it requires stronger assumptions about the character of the interaction.

The basic model is the same as before with the following modifications. Recall that our central purpose is to illustrate the equilibrium when the moderate's position in the later stage is dependent on the amount of suppression in the earlier stage. To explore this type of correlation, we now introduce two types for the in-group. In particular, we assume that any bargain reached in the early stage has some probability of breaking down in the future. This probability, further, depends on the type of the in-group. At the beginning of the game, the probability that M_1 is a high breakdown type is σ. Formally we designate M_1 's type by $\tau \in$

$\boxed{\text{X}}$

(H,L). Further, we assume that while M_1 knows its type, M_2 and R_2 do not.

We also assume that there is a cost for the in-group to suppress, which we denote c_r for the costs of repression. These costs can either be high or low, designated by c_r^H or c_r^L, where $c_r^H > c_r^L$. At the beginning of each turn, we assume that c_r is randomly drawn from a

binomial distribution which is conditional on the in-group's type. Here we assume that if M_1 is a "high type"($\tau = H$), then the probability $c_r = c_r^H$ is ϕ^H, and if M_1 is a "low type"($\tau = L$), the probability $c_r = c_r^H$ is ϕ^L. To capture the notion that low types are more likely to bear high costs, we assume $\phi^L > \phi^H$. Finally, to simplify the analysis, we assume that $c_r^L = 0$ and that c_r^H is very large, say ∞—so when $c_r = c_r^H$, M_1 does not have an incentive to suppress, a bound we will define later.[13] So, in this version of the *TG,* we are assuming that there are some additional costs that cause a disincentive for the in-group to suppress. The fact that these are randomly drawn (although correlated with the type) allows us to introduce incomplete information. In practice, this assumption is reasonable if over time, we assume that there is idiosyncratic error that drives the ease or difficulty in a specific response to terrorism. While the assumption of independence across periods implicit in the above specification are fairly strong, it provides, in our view, a good first cut for the examination of these issues.

The final modification we make in this extension is to the *CWG*. Here we assume that if a bargain is struck, it provides an ongoing payoff. If the bargain is struck between M_1 and M_2 we assume that there is some probability that the bargain will break down in the future— that M_1 will renege, and implement her own preferred policy. Again, here, we assume that the probability of a breakdown in any period is a function of the in-group's type, so that a high type will renege in any turn with probability θ^H and a low type will renege with probability θ^L, where $\theta^H > \theta^L$.

Given this setup, we start by analyzing the *CWG*. Here, the same logic as before applies in that if M_2 rejects M_1 's offer b_1 then M_2's expected payoff is a lottery between her own ideal point and zero. This means, M_2 's average payoff is $-\pi x^2$. Following the same argument as before, we can also write down M_2 's payoff if she accepts b_1. To do this, however, we have to know what M_2 's subjective beliefs about the likelihood of a breakdown are; in other words, given the history of the game to that point, what is the probability that M_1

---

[13] In fact, this upper bound could be much smaller.

is a high type? Since we will define this as part of the equilibrium later, for now we define this probability to be β, which is a function of the history of play. Here, we have her expected payoff being

$$\sum_{t=0}^{\infty} \theta_L^t \beta (b_1 - x)^2 + \theta_H^t (1-\beta)(b_1 - x)^2 + (1-\theta_L^t)\beta x^2 + (1-\theta_H^t)(1-\beta)x^2 .$$

This in turn implies that

$$\pi x^2 = [\frac{\beta}{1-\theta_L} + \frac{1-\beta}{1-\theta_H}](b_1 - x)^2 + x^2 . \tag{1}$$

Solving (1) by the quadratic formula and using a similar maximization argument as appears in the original proof, we have:

$$b_1^* = x(1 - \sqrt{1 - \frac{1-\pi}{\frac{\beta}{1-\theta_L} + \frac{1-\beta}{1-\theta_H}}}) . \tag{2}$$

$\boxed{\times}$

Notice that the result in Proposition 1B of the previous version is simply a special case of this result (when $\theta_L = \theta_H = 0$). A couple of important properties of this result are worth noting. First, as before, the final policy outcome is increasing in *x*. Notice here, however, *x* is fixed and does not move. Second, taking the first derivative of (2) shows that the offer is also increasing in β, the posterior probability that the in-group is a high or "bad" type. In this case, then, the in-group will have an incentive to try to maintain its perception as a low type. Third, taking the second derivative with respect to the posterior probability shows that the magnitude of increase in the offer is diminishing with respect to β. Putting these together shows that *if suppression acts in the TG to increase* β, then the final expected outcome for

the in-group and the out-group will operate in an identical way to the reduced form in the earlier version.

Using this result, we can now analyze the modified version of the *TG*. To solve this game we characterize a Perfect Bayesian Equilibrium (PBE) (not necessarily unique):

Proposition 2. *The following constitutes a PBE to the modified TG.*

(i) M_1 *suppresses when* $c_r = c_r^L$ *if* c_t *is sufficiently high and* β *is sufficiently low. Otherwise it does not suppress.*

(ii) R_2 *terrorize if*
(a) c_t *is sufficiently low and* M_1 *is suppressing*
(b) *or* M_2 *is not suppressing*

(iii) M_2 *updates its beliefs about* M_1 *'s type based on Bayes Rule.*

The proof, which we outline here, provides an intuition behind the results. Note first that by construction, M_1 will not suppress whenever $c_r = c_r^H$. Now consider M_2's and R_2's beliefs about τ. Based on the equilibrium, if M_1 suppresses, then

$$\Pr(\tau = H \mid \sup pression) = \beta' = \frac{\Pr ob(\sup pression \mid H) \Pr ob(H)}{\Pr ob(\sup pression \mid H) \Pr ob(H) + \Pr ob(\sup pression \mid L) \Pr ob(L)}$$

$$= \frac{(1 - \phi^H)\beta}{(1 - \phi^H)\beta + (1 - \phi^L)(1 - \beta)} \tag{3}$$

Notice that this probability is ineasing when M_1 suppresses. To see this, it is sufficient to show that

$$\frac{(1 - \phi^H)\beta}{(1 - \phi^H)\beta + (1 - \phi^L)(1 - \beta)} - \beta > 0$$

Rearranging terms, this condition can be written

$$\phi^L > \phi^H$$

which is true by assumption. In terms of β, we can further analyze the marginal effects with respect to the posterior beliefs. Taking the first and second derivatives of (3), we have

$$\frac{\partial \beta'}{\partial \beta} = \frac{(1-\phi^H)(1-\phi^L)}{[(1-\phi^H)\beta + (1-\phi^L)(1-\beta)]^2} > 0$$

and

$$\frac{\partial \beta'}{\partial \beta^2} = -\frac{2(1-\phi^H)(1-\phi^L)(\phi^L - \phi^H)}{[(1-\phi^H)\beta + (1-\phi^L)(1-\beta)]^3} < 0$$

which means that the change in the posterior beliefs is decreasing in β. Thus, the more fully specified version of the model has a similar dynamic to that in the earlier version: because *Bayesian updating* of beliefs is a function of the prior beliefs, the amount that the beliefs moves is *decreasing* as the prior probability of a high-type in-group increases.

Next we turn to the best response by R_2 given the equilibrium strategies of the other players. When M_1 is not suppressing, the expected value of a single terrorist act is positive since, as before, $\varepsilon > 0$. When M_1 is suppressing when it realizes a low cost, it is sufficient to analyze R_2's single-period incentive to terrorize. Here, the expected payoff to terrorizing is

$$-((1-\phi^L)(1-\beta)+(1-\phi^H)\beta)[(b_{1\beta'}^* - 1)^2 + c_s + \varepsilon] - (1-((1-\phi^L)(1-\beta)+(1-\phi^H)\beta))((b_{1\beta}^* - 1)^2 + \varepsilon) \ (4)$$

which is simply the probability that M_1 will suppress—drawing a low cost of repression—times the payoff for suppressing plus the probability that M_1 will not suppress times the value if it does not. Note we use the notation $b_{1\beta}^*$ and $b_{1\beta'}^*$ to indicate the optimal offers in the CWG given different beliefs about β —under non-suppression and suppression. If chooses not to terrorize, her expected payoff is simply

$$- (b_{1\beta}^{*} - 1)^{2} \tag{5}$$

Setting (4) greater than (5) implies R_2 will terrorize if and only if

$$c_s < (b_{1\beta}^{*} - 1)^2 - (b_{1\beta'}^{*} - 1)^2 - \frac{\varepsilon}{(1 - \phi^{L})(1 - \beta) + (1 - \phi^{H})\beta} = c_s^{*} \tag{6}$$

The condition in (6) is similar to the one in the original version. In particular, if the costs of suppression are too high, the terrorists will not terrorize. The critical point, however, is to notice that by the previous arguments about β, the difference defined by the first two terms on the right hand side of (6) is decreasing in β, and the third term is also decreasing in β. This implies that when β is sufficiently low, terrorists will terrorize, potentially inviting suppression, but when β is high, terror will cease, as the benefit to radicalizing the (beliefs) of moderates will be minimal. This is precisely the same intuition provided in the first model.

The final step to check when the stated equilibrium can be sustained is to analyze the incentives of the in-group M_I. Here, the incentives are potentially different than in the earlier model as M_I knows that pretending to be a low type is advantageous in the later stages. Here, we perform a similar analysis of the incentives of R_2, checking the single-period incentives to suppress. When R_2 is not terrorizing, clearly there is no incentive to suppress. When R_2 is terrorizing, if $c_r = c_r^{H}$, M_I will not suppress since, by construction, c_r is large. The question to answer then is when would M_I suppress when the costs of doing so are low? Here, the expected utility from suppressing given R_2 has terrorized is

$$-\alpha (b_{1\beta'}^{*})^{2} - (1 - \alpha)[(b_{1\beta'}^{*})^{2} + 2c_t] \tag{7}$$

If M_I does not suppress, its payoff is

$$(b_{1\beta}^{*})^{2} - c_t \tag{8}$$

Setting, (7) greater than (8), we have M_1 will suppress iff

$$c_t > \frac{(b_{1\beta'}^*)^2 - (b_{1\beta}^*)^2}{2\alpha - 1} = c_t^* \tag{9}$$

This means that if the cost of terrorism is sufficiently high, the in-group will suppress, trading the benefit of avoiding future terror for the cost of decreasing its long-term payoff by driving the moderate out-group to a less trusting position.

This result is similar to that in the earlier version. The most important result, however, is the way c_t^* behaves with respect to β. The earlier analysis of the behavior of β indicates that the cutpoint c_t^* is decreasing in β. This implies that if M_1 has suppressed in an earlier period, it will continue to have the same incentive. Alternatively, there is some critical value β^* such that if $\sigma > \beta^*$, suppression and cycles of violence will commence. Otherwise, M_1 will never suppress and the cycle will never begin. This result is analogous to that in the earlier model with induced ideal points. The difference is that the critical operation here is in terms of beliefs rather than preferences: while M_2 's ideal point is unmoving, the final outcome is a function of its beliefs about M_1's type. Substantively, however, the results are the same: when moderates' beliefs are truly moderate, cycles of violence are most likely to commence.

6. Discussion

As we noted at the outset, one requirement of a theory of terrorism is that it explain both cross-sectional and dynamic or time-series variation among cases in which the possibility of terrorism exists. In this paper, we advance a model which is a first-cut attempt to explain such variation. In terms of cross-sectional variation, the model highlights the importance of suppression as an important motivator for setting off cycles of violence. If suppression is not a viable strategy for the dominant group, it vastly reduces the incentive of radical or marginal groups to undertake political violence. What factors are likely, then, to make such violence a possible outcome? In our analysis, two factors are crucial. On the one

hand, at the outset, moderates must be indeed moderate—if they are politically close to the dominant group, the dominant group (e.g. Protestants, Israelis, or the Sinhalese majority) will not be willing to risk that relationship in order to suppress terrorist outbursts. On the other hand, if the moderates (e.g. Catholics, Palestinians and Arabs, or Tamils) are very close to their radical counterparts (e.g. the IRA, Hammas, or Tamil Tigers), then the radicals gain very little by inviting attacks upon themselves. The second factor enters in a similar way. If the possibility of eliminating the terrorists is reasonably small, again the incentives to undertake such an act is also limited, and thus once again, the "value of terrorism" is limited to its emotional or preference-driven component.

In this sense, we can return to the initial explanations of political violence that we posited. The model makes two contributions to the strength of the *provocation hypothesis*. First, it shows how the value of provocation vastly increases the likelihood of *sustained terrorist* activity. Second, it defines the conditions under which such activity will both start *and* stop. When regimes are non-suppressive for example, or when suppression is an ineffective tool, such provocation by radicals will not be effective and therefore, terrorist cycles will not be observed.

The model also shows how this hypothesis interacts with the *revenge hypothesis*. One subtle aspect of the model is that suppression only is valuable when terrorism will continue even in the absence of suppression (in the model, for one more turn). This means that the radicals must always have an incentive to terrorize in their last opportunity, even when not suppressed. This only is rational if there is some benefit to terrorism in and of itself (e.g. $\varepsilon > 0$). If this was not the case, then the dominant group would never suppress terror, and therefore, would allow the game to proceed to the bargaining stage. In our model, a desire to the dominant group is an important factor in determining the equilibrium of political violence. On the other hand, without a political motivation, the model also shows the relative weakness of the revenge hypothesis to explain persistent political violence. The reason is that the opportunity to radicalize moderates at once provides a greater incentive for acts of terrorism *and* a rationale for continuing with such acts.

A third explanation for political violence is that it is intended to *impose costs* upon the dominant group in order to gain concessions from the sitting government. In this case, partly by construction, the model indicates such an explanation is of limited value. In the context of

the model, for example, the *higher* the costs of terrorism, the less likely cycles of violence are to break out. Of course, this begs the question of why radicals do not temper their violence to fall within an acceptable range to a suppressive regime. While the model does not address this legitimate objection, it might be difficult to make such a threat of limited violence credibly.

Finally, as we noted earlier, many observers have commented on the fundamentally *irrational* nature of terrorist acts of political violence. Although within the context of a rationalist model it is impossible to theoretically reject such a claim, the fact that our model generates precise (if not yet wholly testable) predictions about both the outbreak and persistence of terrorism, within the structure of strategies pursued by rational actors, the model provides a first cut at showing the (subset of) rational actions taken within the context of political violence.

APPENDIX: Proofs of Proposition 1a and 1b

Using backwards induction, we prove Proposition 1b first, and then Proposition 1a.
Proof of Proposition 1b. R_2 will accept b_2 iff

$$-(1-\pi)(1-b_2)^2 - \pi(1) > -(1-0)^2 \Rightarrow b_2 > 0$$

That M_2 will offer $b_2 = x$ follows directly from R_2's acceptance function above. M_2 will therefore accept an offer of b_1 iff

$$-\pi(0-x)^2 - (1-\pi)(x-x)^2 < -(b_1-x)^2$$

which implies M_2 will play A iff

$$(1-\pi)x^2 - 2b_1x + b_1^2 < 0$$

Solving this quadratic for b_1, we obtain that M_2 will accept iff

$$b_1 \in (x(1-\sqrt{\pi}), x(1+\sqrt{\pi}))$$

Using this acceptance function, we can write down M_1's expected utility as a function of b_1. In particular,

$$EU_{M_1}(b_1) = \begin{cases} -(1-\pi)x^2 & if & b_1 < x(1-\sqrt{\pi}) \\ & or & b_1 > x(1+\sqrt{\pi}) \\ -b_1^2 & otherwise \end{cases}$$

so

$$b_1^* = \arg\max_{b_1} EU_{M_1}(b_1) = \begin{cases} 0 & if & 1-\pi < (1-\sqrt{\pi})^2 \\ x(1-\sqrt{\pi}) & otherwise \end{cases}$$

Since $\pi \in (0.5, 1)$, this can be reduced to

30

$$b_1^* = x(1 - \sqrt{\pi})$$

Which constitutes the proposition. ■

Lemma 1. *If M_1 is not suppressing, R_2 will terrorize.*

Proof of Lemma 1. If M_1 is not suppressing, R_2 will terrorize if

$$-[1 - x(1 - \sqrt{\pi})]^2 + \varepsilon > -[1 - x(1 - \sqrt{\pi})]^2$$

which is true by $\varepsilon > 0$. ■

Lemma 2. *If M_1 suppresses she will always suppress.*

Proof of Lemma 2. Consider M_1 's single-period payoff to suppressing. Here we have her expected utility is

$$-[(1 - (1 - x)\gamma)(1 - \sqrt{\pi})]^2 - (1 - \alpha)c_t \qquad (10)$$

and for not suppressing is

$$-[x(1 - \sqrt{\pi})]^2 - c_t \qquad (11)$$

Thus, if M_1 suppresses this implies (10) is greater than (11). Solving the inequality, we have

$$c_t' \equiv c_t > \frac{(1 - \sqrt{\pi})^2}{\alpha}\left[(1 - \gamma)^2 + 2x\gamma(1 - \gamma) - x^2(1 - \gamma^2)\right] \qquad (12)$$

Taking the first derivative of (12) with respect to x, we obtain

$$\frac{\partial c_t'}{\partial x} > \frac{(1 - \sqrt{\pi})^2}{\alpha}\left[2\gamma(1 - \gamma) - 2x(1 - \gamma^2)\right] \qquad (13)$$

The Lemma follows from the fact that (13) is negative if $x > \dfrac{\gamma}{1+\gamma}$ which is true by assumption. ∎

Lemma 3. *If M_1 suppresses there exists some point c_s^* such that if $c_s > c_s^*$, R_2 will not terrorize.*

Proof of Lemma 3. Consider R_2 's incentives to terrorize for a single period. Here we have that if M_1 is playing S by Lemmas 1 and 2, R_2 will play T if

$$-[(1-(1-(1-x)\gamma))(1-\sqrt{\pi})]^2 - c_s + \varepsilon < -[1-x(1-\sqrt{\pi})]^2$$

$$\Rightarrow c_s < [1-x(1-\sqrt{\pi})]^2 - [1-(1-(1-x)\gamma)(1-\sqrt{\pi})]^2 + \varepsilon = c_s^*$$

The proposition follows directly from the fact that $\dfrac{\partial c_s^*}{\partial x} < 0$. ∎

Proof of Proposition 1a. M_1 will always play NS if R_2 is playing NT since

$$-[(1-(1-x)\gamma)(1-\sqrt{\pi})]^2 < -[x(1-\sqrt{\pi})]^2$$

If R_2 is playing T, then M_1's expected payoff for playing NS is

$$-(x(1-\sqrt{\pi}))^2 - c_t \tag{14}$$

If M_1 is playing S by Lemma 2, R_2 will play T for t^* rounds in which is has the opportunity, where t^* is the minimum integer that solves

$$c_s > [1-x(1-\sqrt{\pi})]^2 - [1-(1-(1-x)\gamma^{t^*})(1-\sqrt{\pi})]^2 + \varepsilon$$

Using this result, the expected utility to M_1 for playing S, by Lemma 3, can be written

$$-\sum_{t=1}^{t^*}(1-\alpha)^{t-1}\alpha[(1-(1-x)\gamma^t)(1-\sqrt{\pi})]^2 + (1-\alpha)^t c_t$$

which can be rewritten

$$-\alpha(1-\sqrt{\pi})^2\left[\frac{1-(1-\alpha)^{t^*}}{\alpha}-2\frac{\gamma-(1-\alpha)^{t^*}\gamma^{t^*+1}}{1-\gamma(1-\alpha)}+(1-x_0)^2\frac{\gamma^2-(1-\alpha)^{t^*}\gamma^{2(t^*+1)}}{1-(1-\alpha)\gamma^2}\right]+c_t\frac{(1-\alpha)-(1-\alpha)^{t^*+1}}{\alpha}\qquad(15)$$

Given Lemmas 2 and 3, this implies that will suppress iff

$$c_{t0}^* \equiv c_t > \frac{\alpha^2(1-\sqrt{\pi})}{2\alpha-1-(1-\alpha)^{t^*+1}}\left[\frac{1-(1-\alpha)^{t^*}-x_0^2}{\alpha}-2\frac{\gamma-(1-\alpha)^{t^*}\gamma^{t^*+1}}{1-\gamma(1-\alpha)}+(1-x_0)^2\frac{\gamma^2-(1-\alpha)^{t^*}\gamma^{2(t^*+1)}}{1-(1-\alpha)\gamma^2}\right]$$

The remainder of the proposition follows from Lemmas 1 through 3. ∎

References (incomplete)

Bates, Robert H., Rui J. P. de Figueiredo, Jr., and Barry R. Weingast. 1998. "The Politics of Interpretation: Rationality, Culture and Transition." *Politics and Society. Forthcoming.*

Corsi, J. R. 1981. "Terrorism as a Desperate Game." *Journal of Conflict Resolution* 25: 47-85.

de Figueiredo, Rui J. P., Jr., and Barry R. Weingast. 1998. "The Rationality of Fear: Political Opportunism and Ethnic Conflict." *Forthcoming*.

Enders, Walter and Todd Sandler. 1993. "The Effectiveness of Antiterrorism Policies: A Vector-Autoregression-Intervention Analysis." *American Political Science Review* 87 (December): 829-844.

Fearon, James D. 1997. "Commitment Problems and the Spread of Ethnic Conflict." In David Lake and Donald Rothchild, eds., *Containing Fear: The International Spread and Management of Ethnic Conflict.*

Fearon, James D., and David D. Laitin. 1996. "Explaining Interethnic Cooperation." *American Political Science Review* (Dec.) 90: 715-35.

Gibbs, Jack P. 1989. "Conceptualization of Terrorism." *American Sociological Review* 54 (June): 329-340.

Laqueur, Walter. 1987. *The Age of Terrorism*. London: Weidenfeld and Nicolson.

Lapan, Harvey E. and Todd Sandler. 1988. "To Bargain or Not to Bargain: That is the Question." *American Economic Review: Proceedings of the One Hundredth Meeting of the American Economic Association* 78 (May): 16-21.

Livingston, Marius H., ed. 1978. *International Terrorism in the Contemporary World.* Westport, CT: Greenwood Press.

Oots, Kent L. 1986. *A Political Organization Approach to Transanational Terrorism*. Westport, CT: Greenwood Press.

Rapoport, David C. 1984. "Fear and Trembling: Terrorism in Three Religious Traditions." *American Political Science Review* 78 (September): 658-677.

Sandler, Todd, John T. Tschirhart, and Jon Cauley. 1983. "A Theoretical Analysis of Transnational Terrorism" *American Political Science Review* 77 (March): 36-54.

Schmid, Alex P. and Albert J. Jongman. 1988. *Political Terrorism*. Amsterdam: North Holland.

Wilkinson, Paul. 1986. *Terrorism and the Liberal State*. New York: New York University Press.