# Mutual Fear and Civil War[1]


Sylvain Chassang[2] — Princeton University

Gerard Padró i Miquel[3] — Stanford Graduate School of Business


September 17, 2007


[1]Previous versions of this paper were presented at the Political Science Departments of MIT, Harvard, Princeton, Columbia, UC Berkeley and Washington University as well as conflict conferences at Northwestern University and NYU. We are thankful for all the comments received. All remaining errors are, of course, our own.



[2]Assistant Professor. Woodrow Wilson School of Public and International Affairs and Department of Economics. 316 Bendhem Hall. Princeton University. Princeton, NJ 08544. (857) 389 0385. chassang@princeton.edu.

[3]Corresponding Author. Assistant Professor of Political Economy. Stanford Graduate School of Business. 518 Memorial Way. Stanford, CA 94305. (650) 724 0166. gpadro@stanford.edu

**Abstract**

We propose a theory of conflict in which actors balance the opportunity costs of fighting with the fear of being attacked. By mobilizing, an agent foregoes returns to her peacetime economic activity, but she can seize resources and protect herself from an attack. Opportunity costs change with the economic situation, which determines the risk of attack in equilibrium. This theory makes two contributions. First, it predicts that conflict occurs after bad economic shocks. This is supported by the empirical literature on civil war and it is difficult to accommodate using existing models. Second, the theory generates conflict out of mutual fears. This is closely related to the literature on the security dilemma. The model allows for a systematic exploration of the theory in Jervis (1978) in a rational choice framework and it shows that aggressive and security seeking agents can be jointly understood as depending on transient economic circumstances.

## Introduction

The prevalence of civil war and the evidence of its disastrous effects has recently motivated a burgeoning empirical literature. It arises clearly from this series of papers that "per capital income is the single best predictor of a country's odds of civil war outbreak, empirically dominating other factors that one might have expected to do better, such as level of democracy, degree of ethnic or religious diversity or nature of ethnic demography, or level of income inequality."[1] Income per capita affects the likelihood of conflict in two different ways. First, poor countries have a higher propensity to see violence. Second, this violence tends to occur when countries suffer an unfavorable economic shock. Hegre and Sambanis (2006) show that these findings are extremely robust in cross-country regressions. The second point is specifically addressed by Miguel et al (2004) who causally establish that a bad economic shock increases the probability of a civil conflict outbreak in Sub-Saharan Africa. This wealth of scholarly evidence concurs with journalistic accounts that have long related bad economic shocks, such as droughts, to ethnic clashes and violence.[2]

These facts are puzzling when seen through the lens of most formal analyses of conflict. This work tipically follows two different traditions. Political scientists have focused on bargaining models

[1]We use the words of Fearon (2005). Collier and Hoeffler (1998, 2004) and Fearon and Laitin (2003) explicitly make the point that economic considerations dominate all other hypotheses. See Sambanis (2002) for an early review of this literature.

[2]See, for instance, "Somalis clash over scarce water", accessed at http://news.bbc.co.uk/go/pr/fr/-/2/hi/africa/4723008.stm; "Ethiopia: Drought stimulates outbreaks of violence" and "Water clash warning evoked by Kenya climate talks" accessed at http://www.reliefweb.int/rw/rwb.nsf/db900sid/EKOI-6VJ5PR?OpenDocument; "Tribal rustlers turn to dealing in death as the drought ends" accessed at http://www.timesonline.co.uk/tol/news/world/article707535.ece.

in which the decision to fight is an outside option used in the case of bargaining breakdown.[3] This literature is concerned with the presence of fighting, which is inefficient, along the equilibrium path, and it typically does not consider the opportunity cost of fighting or mobilizing.[4] Conversely, economists have developed models that include a resource trade-off between production and coercion but do not explicitly consider a decision to fight.[5] In their canonical formulations, none of these models can account for the relationship between income (and income shocks) and violence. In the bargaining case, a bigger size of the pie increases incentives to fight if the costs are constant. If the destruction caused by fighting is proportional to the size of the pie, then these models predict that violence is not affected by income. In the latter type of models, as returns to production decrease, both the opportunity costs of violence and the amount of grabable resources decrease. Hence they also yield income-neutral results.

We propose a model that makes two contributions to the theoretical literature on civil war. First, it provides an intuitive mechanism that makes sense of the empirical patterns detailed above. Second, it puts contenders in a strategic situation in which mutual fears generate civil war, a possibility raised in work by Posen (1993) and Jervis and Snyder (1999).

As a starting point for our analysis, we note that fighting entails an opportunity cost. Indeed, agents typically have a peace-time occupation that yields some returns and has to be abandoned if they mobilize. Hence, to put it in stark terms, militiamen have to choose between the fruits of the ploughshare and the fruits of the AK-47. Imagine, for example, a group that has to decide whether to till its land or to attack a neighboring group to seize some extra land. When the

[3]See, for instance but by no means exclusively, Fearon (1995), Powell (1996a, 1996b, 1999 and 2004a) and Slantchev (2003). Powell (2002) provides a survey of this literature.

[4]See Slantchev (2004) for a recent exception.

[5]Hirshleifer (1995, 2001), Grossman (1991), and Skaperdas (1992) are some early examples of this literature.

economic situation is bad, for instance as a result of a drought, returns to labor are very low. In contrast, the extra land to be seized retains its value: returns to land may be low today, but once the drought abates, controlling a bigger area will provide bigger crops. Therefore, the opportunity costs of fighting are directly dependent on transient economic circumstances while the returns to fighting are not. Hence, it is natural that, in the presence of a drought, the militamen reach for their weapons and look eagerly at their neighbors possessions.

Now imagine that we model two groups living side by side. They receive a common economic shock, and hence both have the same incentives: in the presence of a drought both groups will try to seize land from each other. This is costly: it entails foregoing production and exerting and suffering violence. Morever, if the groups are evenly matched they are not able to make much territorial gains when both are mobilized. In short, when economic circumstances are dire, conflict is *inevitable* as groups are locked in a Prisoners' Dilemma. They would rather live in peace, but each group knows that if it does not mobilize, the other one will attack it and seize the land under contention. Hence, the only equilibrium in the presence of a really bad economic shock implies fruitless violence.

For better economic circumstances, both peace and war are sustainable. Indeed, the game takes on the structure of a Stag Hunt. Groups have no desire to unilaterally attack due to the higher labor returns. However, they mobilize if they expect their opponents to do so, since they are eager to avoid losing land and suffering ethnic cleansing or subjugation. As Jervis (1978) argued, in a Stag Hunt game the bad equilibrium captures the mutual fears that prevent players from coordinating into peaceful cohabitation. Posen (1993) and Jervis and Snyder (1999) develop this idea to argue that in the anarchical environments of weak states, groups with fundamentally compatible goals may end up locked in violent conflict because of the mutual fear of an attack.[6] However, as Fearon

[6] There is a long intellectual history behind this idea. See Hobbes (2005), Schelling (1960) and

and Laitin (1996) point out, ethnic groups are able to coordinate into peaceful cohabitation in most situations, even in weak states. In terms of the model, as both equilibria coexist, it is difficult to ascertain whether groups will be able to coordinate or not. In what circumstances do mutual fears generate actual conflict?

We show that a realistic equilibrium selection criterion applied to the model provides much sharper predictions and a better understanding of the mechanism behind mutual fears. In particular, we use the risk-dominance criterion established by Harsanyi and Selten (1988). This criterion was developed to deal with situations in which players are trying to second-guess each other's moves. Hence it is particularly appropriate to this context of ethnic conflict.

In the selected equilibrium, players change their attitudes towards violence and conquest as a function of the economic situation. For extremely bad economic circumstances, when conflict is *inevitable*, agents are *aggressive* or *greedy* in the sense that they wish to attack the opponent even if they are guaranteed that no attack is forthcoming. For better economic shocks, but still below a well defined threshold, players are *security-seekers* and find themselves locked in conflict because of mutual fears. Such fears force both players to launch preemptive attacks even in economic circumstances in which war is not *inevitable*. In this situation, both groups end up fighting for (rationally) defensive reasons, in a static counterpart of the Security Dilemma. Indeed, in these circumstances, if any player was able to commit not to attack, violence would be averted. Unfortunately, in a weak state no group is willing to renounce its capacity for war because there is always some positive probability that a really bad economic shock will make war *inevitable* in the future.

This equilibrium provides some additional insights on the forces that exacerbate mutual fears in a Stag Hunt situation. First, richer countries see less violence. Indeed, in richer countries both groups know that returns to labor are generally high. This has two positive consequences: firstly,

Jervis (1976). Posen (1993) first applied this idea to civil wars.

each group has smaller incentives to deviate. Secondly, each group also knows that its opponent has fewer incentives to attack. As a consequence, peace is reinforced and mutual fear abates.

Second, worse payoffs in the case of war support more peaceful equilibria. Finally, higher payoffs in a first strike and worse payoffs in suffering such a first strike worsen coordination. These natural comparative statics are in line with the arguments influentially posited by Jervis (1978) and do not exist if we obviate mutual fears and only examine the cases in which war is *inevitable*. Hence, our game-theoretic approach gives a robust rational choice basis to some existing theoretical arguments in international relations.[7]

The paper is organized as follows. Section 2 presents the model. Section 3 discusses the intuition behind the main forces in the model and analyzes it. Section 4 introduces risk-dominance as an equilibrium selection criterion, applies it to the game, and obtains predictions that are in line with the empirical literature. It also shows that the risk-dominant equilibrium is in fact the unique Nash equilibrium of a slightly modified version of the game. Section 5 discusses in depth the relationship between the present findings and previous theoretical arguments of international relations scholars such as Posen (1993) or Snyder and Jervis (1999). Proofs are in the Appendix.

## The Model

Consider two groups, living side by side for two periods. Each group possesses a unit of land at the beginning of the game. In the first period, each group can decide to devote its efforts to tilling

[7]Kydd (1997, 2005) and Baliga and Sjoström (2004) provide formal interpretations of the security dilemma in models based on the potential existence of aggressive types. In contrast, our approach hinges on the common state of the economy. This allows us to link the economic situation to the possibility that fear of miscoordination leads to war. In any case, we view these two approaches to modelling mutual fears as highly complementary.

its land or it can decide to mobilize for violence. Denote the peaceful strategy $p$ and the violent strategy $a$ (for attack).

If a group works (plays $p$), its returns depend on the productivity of land, $\pi$ and the amount of rainfall, $\theta$. $\pi$ captures the productive resources of a group, such as irrigation, tools, animals or land quality, or more generally anything that makes labor more productive, such as physical capital. Given that we are keeping land holdings equal, $\pi$ is also a measure of wealth. $\theta$ captures a transient economic state of the world that determines returns to labor. In the agricultural example we use, rainfall would cause such fluctuations to returns to labor. More generally, $\theta$ captures forces such as world market prices or investment flows. $\theta$ is common to both groups and is randomly distributed as a uniform $U[\Theta - L, \Theta + L]$, where $\Theta \geq L$. There is a rainfall realization in the first period, called $\theta_1$, and a second different and independent rainfall realization for the second period, $\theta_2$.

A group can decide not to work and instead attack the neighbor. This may allow the group to seize some extra land. Land captured is valuable in the first period because of the returns in the second period. If a group attacks while the neighbor is not mobilized, it can seize $\delta < 1$ units of land from the neighboring group. The group that is attacked loses these $\delta$ units of land and on top of that suffers costs $S$ which capture the violence and atrocities associated with processes of ethnic cleansing and subjugation. Moreover, the attacking group can seize the production that the attacked group generated with its work on the land that is seized.[8] To capture in a stark manner that there is an opportunity cost to fighting, we assume that when a group attacks (plays $a$), it does not produce in its land for that period as men devote their efforts to violence instead of productive activities.

If both groups fight (play $a$), no land changes hands, and both groups pay a cost $W$ of waging

[8]For simplicity, we assume that no production is lost in the attack. This is not essential: the model can easily accommodate additional costs of attacking.

war. We make an assumption on the relationship between these constants.

**Assumption 1:** $S - W > 0$

Intuitively, this condition says that ethnic cleansing is more painful than the fight to avoid it: it ensures that groups have an incentive to fight if they expect an attack. This is important because it introduces a *defensive motive* for fighting: groups may decide to mobilize and abandon their economic activity not because they want to attack the neighbors but because they fear an attack. In fact, in most cases in which the balance of forces is quite even, $S - W$ should be huge as defending could stop a genocide from happening.

The timing of this game is as follows:

1. $\theta_1$ is realized and observed by both groups.
2. Players decide whether to play peace, $p$ or violence $a$.
3. Land changes hands in the case of unilateral attack, $\theta_2$ is realized, and players obtain their payoffs.

With this timing, if neither group fights, each of them receives (in expected value)

$$U\left(p, p; \theta_1\right) = \pi\theta_1 + \beta\pi\Theta$$

where the first term is the current payoff, which is known once the rainfall shock is realized, and the second term is the expected value of a unit of land in the second period (with time discount factor $\beta$). For simplicity, let us call $T \equiv \beta\pi\Theta$, the expected present value of a unit of land.

If one group attacks the other, payoffs are as follows

$$\begin{aligned} U\left(a, p; \theta_1\right) &= \delta\pi\theta_1 + (1+\delta)T \\ U(p, a; \theta_1) &= \left(1-\delta\right)\pi\theta_1 + \left(1-\delta\right)T - S \end{aligned}$$

where the first payoff is the one received by the attacker and the second one is the payoff of the victim. Again, the first additive term is current consumption taking into account that the attacking group seizes current production on the land captured, but it does not produce on the land owned. The second additive term is the expected value of land adjusting by the amount captured (or lost). Finally, if both groups attack, they receive $T - W$ as no production is done in the first period and no land changes hands.

The game can thus be expressed in normal form as a function of the rainfall realization, $\theta_1$. The following matrix states the payoffs for the row player (the payoffs for the column player are symmetric):

| | $p$ | $a$ |
|---|---|---|
| $p$ | $\pi\theta_1 + T$ | $(1-\delta)\,\pi\theta_1 + (1-\delta)\,T - S$ |
| $a$ | $\delta\pi\theta_1 + (1+\delta)T$ | $T - W$ |

(1)

Note that the strategic situation in this model can be applied to settings beyond the one described. An alternative interpretation, also related to many instances of civil war, is one in which two groups are battling for extended control of the state. In this case, in the status quo groups obtain an equal share of the resources generated by the state, $2\pi\theta_1$. The shock $\theta$ is then reinterpreted as tax revenue and foreign aid transfers that can be captured by the contenders. In this case, when a group does successful battle against and opponent, it captures an extra slice of the state of size $\delta$. For expositional ease, we will keep using the land metaphor throughout, but we will return to the control of the state when we discuss the applicability of the results.

## Analysis

**Intuition** Before formally solving the game, let us analyze the strategic forces at play. When choosing whether to work or to attack, players face different motives as a function of their expectations about the actions of their opponents.

If a player expects her opponent to be peaceful, her decision is mostly a function of comparing the potential gains with the opportunity cost of fighting. If she decides to attack, she forgoes the fruits of work in her own land but gains the *future* fuits of enlarged landholdings. The current returns to work depend on the current realization of rainfall. In contrast, the value of holding more land does not typically depend on current, transient economic conditions. Hence, it is intuitive that players will decide to attack when current rainfall shocks are bad: gains are independent of current conditions but opportunity costs are necessarily linked to current returns to labor.

When the enemy is expected to attack, the decision whether to attack or not is given by defensive imperatives provided by Assumption 1: by reverting to violence in the face of an attack, a group gains in two ways. First, it avoids the seizure of its lands. Second, it avoids the suffering associated with ethnic cleansing or subjugation. Hence, for most rainfall realizations, the expectation of an attack will be met by violence.

This combination of incentives predicts that violence should flare for low returns to labor (low realizations of $\theta_1$), while peace should be sustainable when the economic circumstances are good due to the implied opportunity cost of fighting. However, the fact that there is also a defensive motive for fighting, points to the possibility of a coordination problem as best responses vary with the expected action of the opponent. Next, we turn to the formal analysis of the game to confirm the intuition laid down here.

**Equilibrium Behavior** To begin the analysis, it is important to realize an implication of Assumption 1:

**Lemma 1** *If Assumption 1 is maintained, there are no $\theta_1$ realizations for which a strategy profile $(a, p)$ or $(p, a)$ can constitute a Pure Strategy Nash Equilibrium.*

This lemma states that the defensive motive is important enough so that for the realizations of $\theta_1$ that impel a group to attack, the best response is to also resort to violence. In other words, when Assumption 1 is maintained, there can be no equilibrium in which a group passively accepts ethnic cleansing and the seizure of their lands.

Recall that $\theta_1$ is realized and perfectly observed by both players previous to the simultaneous decision to attack. The following proposition states the set of equilibria in the game as a function of the realized $\theta_1$.

**Proposition 1** *There are thresholds $\underline{\theta}$ and $\bar{\theta}$ such that the game displays the following three regimes as a function of the transient economic circumstances, $\theta_1$.*

1. *When times are really bad, $\theta_1 < \underline{\theta}$, there is a unique Nash equilibrium at $(a, a)$.*
2. *For intermediate economic circumstances, $\underline{\theta} < \theta_1 < \bar{\theta}$, there are multiple Nash equilibria at $(p, p)$ and $(a, a)$.*
3. *When times are very good, $\theta_1 > \bar{\theta}$, there is a unique Nash equilibrium at $(p, p)$.*

*where $\underline{\theta} = \frac{\delta T}{\pi(1-\delta)}$ and $\bar{\theta} = \frac{\delta T}{\pi(1-\delta)} + \frac{S-W}{\pi(1-\delta)}$*

Proposition 1 is easy to verify with a few calculations. First, note that when $\theta_1 < \underline{\theta}$ the best response for a group expecting their neighbors to be peaceful is to attack them in order to seize land. To see this note that the prize obtained in case of deviation equals $\delta T$, a constant. Conversely, the opportunity cost, taking into account that some production is seized with the land equals $(1-\delta)\,\pi\theta_1$ and is increasing in the realization of rainfall. It is clear that for $\theta_1$ low enough this opportunity cost will be below the prize. This relationship determines $\underline{\theta}$. Unfortunately, at this $\theta_1$ the other group also decides to attack. As a consequence, the only equilibrium possible is one of generalized violence. Essentially, when $\theta < \underline{\theta}$ the game has the same structure as a Prisoners'

Dilemma. The cost of this fighting is double. First, groups forgo the (meager) returns to work in the first period. Second, they bear the cost of fighting $W$. However, the opportunity cost of fighting is too small and as a consequence they could not commit not to attack even if they expected the opponent to play peace. In these situations, we say that conflict is *inevitable.*

In the other extreme, when $\theta_1 > \bar{\theta}$ the productivity of the land is so high that a group will decide to play peace and work even if it expects an attack. In this case, returns to labor are so high that they trump the defensive motive for fighting. This makes peace a dominant strategy and ensures that no player does, in fact, entertain the possibility of fighting. Hence in this case peace is the only equilibrium.

In general, these two extreme cases above happen quite infrequently. Hence, we can expect that for most realizations of $\theta_1$, we are in the intermediate regime in Propostion 1. In this range, $\theta_1 \in \left(\underline{\theta}, \bar{\theta}\right)$, the concurrence of the opportunity cost of fighting and the defensive motive for fighting generates multiple equilibria and both peace and war are sustainable. Why is this the case? On the one hand, if a player expects her opponent to play peace, she has no reason to deviate and attack because the opportunity cost is relatively high. In this case, the peace equilibrium would be played. On the other hand, if a player expects an attack, she will respond with mobilization and violence because in most normal economic situations the defensive motive is stronger than the opportunity cost of fighting. As a consequence, it is easy to verify that for these intermediate values of $\theta_1$, the normal form of the game displays a Stag Hunt structure. Figure 1 provides an illustration of this equilibrium structure.

Insert Figure 1 about here

It is apparent from Proposition 1 that this model exhibits a dependence of conflict on economic circumstances that is in line with the empirical record of civil war. For disastrous economic situations, conflict is *inevitable.* For better returns to labor, conflict may, or may not, happen depending on the ability of groups to coordinate into peace. Finally, when the economic situation is very buoyant, there is no violence. Therefore in this model a bad economic shock dramatically increases the likelihood of civil war. Previous models that take into account the opportunity cost of fighting by having agents devote resources to coercive activities have a fully static nature.[9] As a consequence, if one reduces the total amount of resources in the economy, both the grabable resources and the returns to labor go down at the same time and no clear relationship between income and the likelihood of violence arises.[10] The model we propose solves this problem by separating the current state of the economy from the future value of productive assets seized. Temporarily low returns to labor provide a perfect window of opportunity to abandon one's economic activity to capture some assets that will be productive with high probability in the future. As Proposition 1 shows, this mechanism can explain the relationship between bad economic shocks and the prevalence of violence.

**Multiplicity and Predictive Power** In an extremely influential article, Jervis (1978) proposed the Stag Hunt game as a framework to analyze anarchy and mutual fears and to base his discourse on which forces ameliorate risk in these situations. We have seen in the previous subsection that as the economic situation improves, the strategic characteristics of the game naturally transition from a Prisoner's Dilemma to a Stag Hunt. Hence it is apparent that in our model of Civil War it is also entirely possible that mutual fears generate conflict in situations in which it was not *inevitable*. In particular, in all the intermediate zone the existence of the $(a, a)$ Nash Equilibrium reveals that

[9]See, for instance Grossman (1991), Hirshleifer (1999) or Skaperdas (1992).

[10]Dal Bó and Dal Bó (2006) present a general equilibrium model in which a similar mechanism is at play. A shock that increases returns to labor will reduce conflict by increasing the opportunity cost of fighting. However, in that model a shock that increases returns to capital will increase violence as the amount of appropriable resources increases.

conflict can occur when a much better option was available.

Note that the width of this middle interval is proportional to $S - W$. If mobilizing for violence does not (also) serve a *defensive motive*, there is no coordination problem and therefore no multiplicity or equilibria. The bigger are the defensive gains, the wider is the area in which this ambiguity exists. This accords with the intuition that the bigger the losses of being caught by an unprepared attack, the higher is the power of mutual fears to generate conflict.

Note, however, that the very fact that allows us to talk about mutual fears generating violence, namely the coexistence of a peaceful equilibrium with a violent one, makes actual predictions difficult. Indeed, as long as both equilibria exist it is difficult to determine which of the outcomes will actually take place. If the peaceful equilibrium exists, what precludes players from coordinating? In short, as long as we have multiple equilibria, game theoretic predictions are difficult to square with the the insightful discussion of Jervis (1978).

In fact, in the model proposed here one can generate a huge number of equilibria as a function of $\theta_1$ as long as the structure in Proposition 1 is respected. Figure 2 illustrates an equilibrium in which players switch from peace to war and viceversa at several points $\alpha_i \in \left(\underline{\theta}, \bar{\theta}\right)$.

Insert Figure 2 about here

This equilibrium might be considered unnatural because of its non-monotonic structure: war occurs for some circumstances that are better than others that support peace. Hence, one might want to consider only a natural equilibrium structure defined by a threshold $t$ in which the equilibrium strategies are to play attack, $(a, a)$ for $\theta_1 < t$ and peace, $(p, p)$ for $\theta_1 > t$. We still have the problem that any $t \in \left[\underline{\theta}, \bar{\theta}\right]$ can constitute an equilibrium of the game. Absent a theory of equilibrium selection, any of these equilibria is equaly plausible. If the criterion is payoff efficiency, then the equilibrium defined by $t = \underline{\theta}$ would be the best one: players are able to coordinate into playing peace for every realization of $\theta_1$ in which peace is sustainable as an equilibrium. Note however, that in this equilibrium mutual fears play no role at all and, as a consequence, many of the insights of Jervis (1978) do not apply: in particular, neither $S$ nor $W$ appear in the expression for $\underline{\theta}$. Can we find an established and intuitive equilibrium selection criterion that allows us to sharpen the predictive power of the mutual fears hypotheses?

## Risk Dominance and Civil War

In this section we propose the risk dominance criterion as a way to select an equilibrium that captures the notion of mutual fears. First we describe this criterion with an example. Second, we apply it to our model and relate the insights obtained to the arguments in Jervis (1978). Next, following Carlsson and Van Damme (1993), we show that this equilibrium is the unique equilibrium of a modified version of the game that is both realistic and appealing in the civil war context we are interested in.

**Risk Dominance** It is convenient to illustrate the concept of risk dominance by way of example. Consider the following 2-by-2 game:

| | $L$ | $R$ |
|---|---|---|
| $U$ | $4, 4$ | $-100, 0$ |
| $D$ | $0, -100$ | $0, 0$ |

where the first payoff in each cell corresponds to the row player. It is clear that this game features two strict Pure Strategy Nash Equilibria at $(U, L)$ and $(D, R)$ and displays a Stag Hunt structure. The first equilibrium Pareto Dominates the second and hence both players would like to coordinate into playing it. However, there is an intuitive sense in which, for the row player, playing $U$ seems more risky than playing $D$. By playing $D$, the row player ensures herself of a payoff of at least 0, while $U$ may end up yielding $-100$. In theory, this payoff should not matter as by using the concept of Pure Strategy Nash Equilibrium, we assume that players play best responses to each other. In practice, however, players need to be wary of ending up in such a situation. Maybe the opponent does not really understand the game. Or maybe she is also nervous about the possibility of the first player deviating. This same nervousness is emphasized by Jervis (1978) in his discussion of the Stag Hunt. In fact, experimental studies in coordination games clearly show that these types of risks matter.[11]

Harsanyi and Selten (1988) systematized this intuition and defined the risk dominance criterion to select among equilibria. This criterion establishes that the equilibrium with the highest product of deviation losses risk dominates all other equilibria.[12] What does this mean in practice?

In the example matrix above, it is easy to calculate deviation losses. Take first the $(U, L)$ equilibrium. If the row player deviates from this cell and plays $D$ instead of $U$, she loses 4 utils. Equally, if the column players chooses $R$ instead of $L$, she loses 4 utils. Hence, the product of deviation losses from the $(U, L)$ equilibrium is 16. However, the deviation losses from $(D, R)$ are much bigger. If the row player decides to play $U$ instead of $D$, her loss is 100 utils. The same is true for the column player. Therefore the product of deviation losses from $(D, R)$ is 10000! Since 10000 is much higher than 16, Harsanyi and Selten conclude that $(D, R)$ risk dominates $(U, L)$ and according to this criterion we should expect it to be played more often.

In Appendix B we discuss the justification that Harsanyi and Selten provide for this criterion, as applied to the civil war payoff matrix (1). The barebones intuition is, however, quite clear. Let

[11]Cooper et al (1990) show that pareto superior equilibria are not necessarily chosen. Battalio et al (2001) show that the risk dominant equilibrium is likely to emerge if best-response calculations are difficult.

[12]A deviation loss from an equilibrium is the loss in utility that a player incurs by changing her action, keeping the actions of the other players constant.

us call the row player $A$ and the column player $B$. When $A$ looks at the matrix and decides what to play, she needs an assessment of what $B$ will do. For every equilibrium, $A$ is worried about two things. First, she needs to determine whether $B$ will deviate. Second, she needs to determine whether $B$ thinks that $A$ will deviate. When she looks at equilibrium $(U, L)$, she realizes that by deviating $B$ only loses 4 utils. In addition, $A$ knows that $B$ is trying to second-guess her at the same time. Therefore $A$ knows that $B$ is also realizing that $A$ would only lose 4 utils by deviating. Hence $A$ concludes that the equilibrium $(U, L)$ is quite risky to be played because there is a high temptation to deviate on *both* sides. In contrast, when $A$ looks at the $(D, R)$ equilibrium, she realizes that $B$ will neither deviate (she would lose 100) nor be worried that $A$ could deviate (she knows that $A$ would lose 100 utils in that case).

It is thus intuitive that the relevant risk is the compounding effect of the fear of a deviation by $B$ and the fear of $B$ expecting a deviation from $A$. This intuition is mathematically captured by the product of deviation losses. The equilibrium that features the highest product of deviation losses is more secure because players expect less deviations from their opponents, and moreover players expect their opponents to be more trustful. The risk dominance criterion thus captures the fact that players have to second guess their opponents and know that their opponents are trying to do the same.

This criterion is especially appropriate for the forces of mutual fear that Jervis (1978) emphasized in the Stag Hunt. Note that the second and the third concerns in the following quote from the aforementioned article are directly related to the risk dominance criterion described above.

> "Unless each person thinks that the others will cooperate, he himself will not. And why might he fear that any other person would do something that would sacrifice his own first choice? The other might not understand the situation, or might not be able to control his impulses if he saw a rabbit, or might fear that some other member of the

group is unreliable."

What do we obtain when we calculate deviation losses from equilibrium in the Civil War matrix (1) that we proposed? Given $\theta_1$, the deviation loss from the peace equilibrium $(p, p)$ equals

$$\pi\theta_1 + T - (\delta\pi\theta_1 + (1+\delta)T) = (1-\delta)\,\pi\theta_1 - \delta T$$

Since $\delta < 1$, the losses when deviating from the peaceful equilibrium are increasing in the returns to labor $\pi\theta_1$. This is in line with the intuition emphasized in Section 3: the higher are returns to labor, the bigger are the opportunity costs of attacking. In the same way, the deviation loss from the violent equilibrium $(a, a)$ can be calculated as

$$T - W - ((1-\delta)\,\pi\theta_1 + (1-\delta)\,T - S) = S - \delta T - W - (1-\delta)\,\pi\theta_1$$

With these calculations, we are ready to apply the risk dominance criterion established by Harsanyi and Selten (1988). The peaceful equilibrium $(p, p)$ risk dominates the conflict equilibrium $(a, a)$ whenever the product of deviations is larger. In symbols, this occurs whenever

$$((1-\delta)\,\pi\theta_1 - \delta T)\,((1-\delta)\,\pi\theta_1 - \delta T) > (S - \delta T - W - (1-\delta)\,\pi\theta_1)\,(S - \delta T - W - (1-\delta)\,\pi\theta_1) \tag{2}$$

Note that this inequality can switch as a function of the realized state of the economy, $\theta_1$. As mentioned above, the left hand side (deviations from peace) is increasing in $\theta_1$ while the right hand side (deviations from war) is decreasing in $\theta_1$. It follows that there is a single threshold $\theta^{RD}$ in which (2) holds with equality. Hence, above the threshold, peace is risk dominant and below it, violence is risk dominant. This threshold equals:

$$\theta^{RD} = \frac{\delta T}{\pi(1-\delta)} + \frac{S - W}{2(1-\delta)\pi} \tag{3}$$

This discussion establishes the following proposition:

**Proposition 2** *Applying the risk dominant criterion to the game selects a unique equilibrium. This equilibrium has a threshold form. For $\theta_1 < \theta^{RD}$, the violent equilibrium $(a, a)$ is risk dominant. For $\theta_1 > \theta^{RD}$, the peaceful equilibrium $(p, p)$ is risk dominant.*

**What Makes Peace More Likely?** Having established that the risk-dominant criterion selects a unique threshold equilibrium, we need to determine whether this equilibrium can make sense of the empirical patterns observed in the data and whether it conforms to the intuition of Jervis (1978) as to when players find it easier to coordinate into peace.

According to Proposition 2, this equilibrium has a very simple structure. As it is shown in Figure 3, the equilibrium is defined by a single threshold, $\theta^{RD}$. For labor returns above $\theta^{RD}$, players play $p$. For returns below, the opportunity cost of fighting is too small and they play $a$.

Insert Figure 3 about here

The risk dominant threshold displays many features that are compelling.

First and foremost, note that fighting happens in the presence of a bad economic shock. The risk dominance selection criterion selects a unique equilibrium that implies fighting for bad economic circumstances and peace for better ones and hence it reinforces the findings of Section 3. As discussed in the introduction, it is an established fact of the cross-country empirical literature on civil war that bad economic shocks generate violence.

These cross-country findings on conflict are also corroborated by some recent research that specifically targets the link between opportunity costs and violent action within countries. Hidalgo et. al (2007) show that land invasions in Brasil happen immediately after adverse economic shocks. Similarly, Dube and Vargas (2007) link violent actions in Colombia with low opportunity costs of agricultural labor by using crop prices.

There is also ethnographic evidence on conflict in anarchical premodern societies that most closely resemble a Hobbesian state of nature. In a survey, Snyder (2002) contends that "research based on the Standard Cross-Cultural Sample suggests that the experience of an unpredictable ecological disaster leading to extreme food shortage is the strongest predictor of an increased likelihood of war." This evidence supports the idea that, in a situation of anarchy, bad economic circumstances help explain the prevalence of conflict.[13]

Case studies on modern civil wars also confirm the relevance of the model. For instance, Flint and De Waal (2005) and Prunier (2005) attribute part of the causes of the war in Darfur to the sustained drought. Also, as discussed in section 2, the model can alternatively be interpreted as two factions fighting for control of the state. In this case, the economic shock $\theta$ models the amount of aid and taxes that can be appropriated by the party that controls it. The equilibrium would predict that fighting occurs with higher likelihood when foreign aid dries out. Laitin (1999), for instance, describes how fighting in Somalia and other African countries became fiercer as foreign aid flows diminished at the end of the Cold War. The same process is described in Rwanda by Prunier (1995).

Second, in the risk dominant equilibrium there is a range of realizations of $\theta_1$ in which conflict happens when it is in fact not *inevitable*. Note that, given Assumption 1, it is always the case that $\underline{\theta} < \theta^{RD} < \bar{\theta}$. In fact, it is immediate to see that $\theta^{RD} = \underline{\theta} + \frac{S-W}{2(1-\delta)\pi}$. For $\theta_1 \in \left(\underline{\theta}, \theta^{RD}\right)$, mutual aprehension predicts the $(a,a)$ equilibrium in a situation in which the $(p,p)$ equilibrium is sustainable. Therefore, the model provides a framework to understand how technological and environmental conditions exacerbate mutual fears.

The following proposition states the comparative statics of $\theta^{RD}$.

**Proposition 3** *The risk dominant threshold, $\theta^{RD}$ is:*

- *Increasing in $S$, the utility loss in case of an undefended attack and $\delta$, the proportion of land*

[13] See, in particular, Ember and Ember (1992, 1997).

*grabbed*

- *Decreasing in $W$, the utility loss in case of a war and $\pi$, the productivity of land*

According to the model and the risk dominant criterion, the probability that $\theta_1 < \theta^{RD}$ provides the likelihood of violence in a given country. Hence, the probability that a country will see internal violence is increasing in $S$ and $\delta$ and decreasing in $W$ and $\pi$, for a given distribution of $\theta$.

Note that these comparative statics are perfectly in line with the hypotheses of Jervis (1978). He writes that the changes of cooperation (peaceful cohabitation) in the context of a Stag Hunt game will be increased by

> "(1) anything that increases the incentives to cooperate by increasing the gains of mutual cooperation and/or decreasing the costs the actor will pay if he cooperates and the other does not; (2) anything that decreases the incentives for defecting by decreasing the gains of taking advantage of the other and/or increasing the costs of mutual noncooperation; (3) anything that increases each side's expectation that the other will cooperate."

These hypotheses are supported by the risk dominant equilibrium of the civil war game.

First, note that $S$, the costs of suffering a genocide, captures the costs of cooperating when the other actor does not cooperate. Hence, increasing $S$ makes cooperation more difficult and increases the likelihood of violence, as Jervis (1978) predicts in point (1). When the costs of ethnic cleansing and subjugation are high, $\theta^{RD}$ increases for two reasons: given the same chance of being attacked, players decide to respond more violently. Moreover, it becomes obvious to both players that in fact, the chances of being attacked actually increase as the other player is using the same reasoning. This last point emphasized in the intuition behind the risk dominance criterion, corresponds to point

(3) in the quote above. It is in fact very natural that groups mobilize for violence more often when the costs of being caught unaware are very high.[14]

Second, not suprisingly, the bigger are the gains from a first strike $\delta$, the more difficult it is for players to coordinate into peace. This is in line with point (2) in the quote above. In the model, this is true for two reasons: for starters, the straight opportunity cost motive becomes more problematic as returns from attacking go up. This can be seen by realizing that $\frac{\partial \theta}{\partial \delta} > 0$. On top of that, increasing $\delta$ magnifies the defensive component of $\theta^{RD}$ because in addition to avoiding the costs of ethnic cleansing net of the cost of fighting, $S - W$, mobilizing for violence also precludes the opponent from seizing $\delta$ units of land.[15]

Third, the risk dominant equilibrium features a characteristic that is reminiscent of deterrence: societies with higher $W$ are more peaceful. $W$ captures the costs of an armed confrontation when both sides are prepared for it. Again, Jervis (1978), in point (2) above states this result. $W$ here clearly increases the costs of mutual noncooperation.

Finally, note that the unique equilibrium selected by risk dominance displays a feature that is in line with most empirical research: poorer societies fight more often. In our model, wealth is captured by $\pi$, as richer societies typically have much higher returns to labor. To understand this comparative statics, note that $\underline{\theta}$ is invariant with respect to $\pi$ because $T$ is proportional to $\pi$. In contrast, the costs of suffering a genocide or the costs of subjugation are not proportional to the marginal product of labor and hence the second additive term in (3) is diminishing in $\pi$. As a result, when players balance their returns from the economic activity with the risk of being caught unguarded, they tend to take more risks as the economy is richer. Higher $\pi$ also means that it is known that the opposite group has higher opportunity costs of fighting. This again clearly helps

[14]See De Figueiredo and Weingast (1999).

[15]This second point can be noted in the second additive element of equation (3).

groups escape the trap of mutual fears.[16]

It is important to note that these natural comparative statics (with the exception of $\delta$) do not apply to the efficient equilibrium, which has the threshold at $\underline{\theta}$. Therefore, the risk-dominant criterion selects an equilibrium that is much more aligned with the theory and intuition of mutual fears as described by Jervis (1978) than the efficient criterion usually invoked in applied formal theory.

A question that one might ask is whether economic inequality across groups exacerbates or mitigates the effects of mutual fears. In the appendix, we develop a variant of the model that allows for inequality in $\pi$ and we denote by $\Delta = (\pi_A - \pi_B)$ the difference in productivity across groups. We obtain the following result:

**Proposition 4** *The risk dominant threshold is increasing in inequality. In symbols,* $\frac{\partial \theta^{RD}}{\partial \Delta} > 0$ *at* $\Delta = 0$*.*

The cross-country empirical literature does not typically find a link between inequality and civil war onset (see, for instance, Fearon and Laitin 2003 or Collier and Hoeffler 2004). However, these articles do not measure inequality across groups but inequality in the country as a whole. The case studies contained in Collier and Sambanis (2005a, 2005b) suggest that inequality, or perceptions of inequality across groups play a role in the onset of violent conflict. Moreover, in within country studies, Hidalgo et al (2007) and Dube and Vargas (2007) find that inequality is related to civil conflict. More empirical work is needed to establish whether inequality across groups is indeed a

[16]Fearon and Laitin (2003) argue that the relationship between income and violence is due to the weakness of the state in poor countries. Fearon (2005) provides a model of this. We are interested in the coordination aspect of violence and therefore we do not explicitly model the state. We view the two forces as complements rather than substitutes.

cause of war. In any case, we show in Proposition 4 that inequality in the productivity of labor, which is typically directly linked to the stock of productive capital, exacerbates the coordination problem of insecure groups and makes mutual fears more conducive to violence.

To summarize, this section shows that the risk dominance criterion for equilibrium selection has many characteristics that are desirable for analizing civil war and the destabilizing effects of mutual fears. On the one hand, its predictions are in line with the empirical literature on civil war. On the other hand it provides a rational choice foundation to intuitions about cooperation that have been in the literature at least since Jervis (1978). There is in fact a third reason that makes risk dominance desirable as a selection criterion, and we expose it next.

**Global Games and Risk Dominance** A slight and realistic modification of the baseline game strengthens the foundation for using the risk dominance criterion. Keep the same payoff structure of the game and imagine that players do not observe $\theta_1$ directly, but observe a (very) precise signal about it. In particular, assume that they observe a signal $s_i = \theta_1 + \varepsilon_i$ where $\varepsilon_i$ is uniformly distributed $U[-\frac{1}{2}\epsilon, \frac{1}{2}\epsilon]$ and $\varepsilon_i$ and $\varepsilon_j$ are independent. This is modelling the fact that players observe the same environment but cannot be sure whether the other player is extracting the same exact conclusions they are. Hence, for small $\epsilon$, their assessment of the state of the world will be very correlated but still not identical. We denote by $\Gamma_\epsilon$ the resulting game with incomplete information.

This information structure corresponds to the framework of global games developed by Carlsson and van Damme (1993). They show that as $\epsilon$ goes to 0, the set of rationalizable strategies converges to the risk dominant equilibrium. Formally, we state the result in Proposition 5.

**Proposition 5** *Assume that $Pr(\theta_1 < \underline{\theta}) > 0$ and $Pr(\theta_1 > \bar{\theta}) > 0$. Then, as $\epsilon$ goes to 0, the set of rationalizable strategies of game $\Gamma_\epsilon$ converges to a singleton constituted by the risk dominant strategies.*

This means that in the slightly modified version of the game, where risks are made explicit by the possibility that observations of the state of the world are different, there is a unique equilibrium that corresponds to the risk dominant equilibrium of the original game. The application of global games to our context is particularly appealing: when a group looks at the sky and predicts whether rain will be forthcoming or not it cannot be sure of how many clouds the neighboring group is noticing. This makes the risk of miscoordinating real. In such circumstances, the group will balance the returns to peace with its assessment of the propensity of the opponent to attack.

Appendix C provides a discussion of the mathematical intuition behind this result. It is instructive, however, to discuss why a threshold at $\underline{\theta}$, which would ensure efficiency, cannot be an equilibrium in this modified game.

Suppose that both players try to coordinate in a threshold equilibrium at $\underline{\theta}$. Hence, they are supposed to play peace if their *private* signal is above $\underline{\theta}$ and attack only if it is below. Note, however, that they do not know the signal that their opponent is receiving. Indeed, their opponent might receive a signal slightly above, or slightly below their own. The noise structure ensures that whatever her current signal $s_i$, player $i$ believes the opponent has a worse (lower) signal with a 50% chance. Now consider the problem of a player that receives a signal exactly equal to $\underline{\theta}$. Given the strategies they are supposed to follow, she knows that she will be attacked with probability 50% as this is the probability that the opponent will get a signal below $\underline{\theta}$. It is easy to check that at $\theta_1 = \underline{\theta}$, with a 50% chance of being attacked the best response is actually to attack: the opportunity costs of labor are too low to risk such high probability of violence. Hence it is clear that no threshold strategy at $\underline{\theta}$ can be an equilibrium of this game. Indeed, it turns out that the threshold at which players are indifferent between attacking or producing given a 50% chance of being attacked is $\theta^{RD}$. This is the reason why the unique equilibrium of this game is the risk dominant equilibrium of the original one.

## Civil War and the Security Dilemma

Starting with Herz (1950) and Schelling (1960) and structured in Jervis (1976, 1978) there is a realization of the fact that the decision whether to fight or not can be seen as a coordination problem. If there is any uncertainty on whether the opponent will attack, chances are that players respond by attacking (or mobilizing, which also wastes resources) to avoid being caught unprepared. This situation is at the heart of the concept of the security dilemma and it is a staple theory of realist security studies.[17] In an influential article, Posen (1993) pioneered the application of the concept of the security dilemma to the analysis of the civil wars that were seemingly burgeoning at the end of the cold war. In its strict version, the security dilemma is a dynamic concept: a security-seeking state, sensing that its neighbors might have predatory inclinations, decides to arm. This makes neighboring states believe that this state might want to attack and they react with more arming. This pernicious spiral is expected to lead to war. Clearly, the model presented here, with its static nature, is not a good depiction of the spiral version of the security dilemma.

However, Jervis (1978) presents a discussion of anarchy and mutual fears in terms of the Stag Hunt, a static game. Equally, in its application to the civil war problem, Jervis and Snyder (1999) define the security dilemma as a "situation in which each party's efforts to increase its own security reduce de security of the others. This situation occurs when geographical, technological, or other strategic conditions render aggression the most advantageous form of self-defense." This depiction of the security dilemma is much closer to the mutual fears mechanism that the model illuminates. For this mechanism to be applicable one only needs two conditions that were already described in the original work of Jervis (1976): an anarchic environment and indistiguishability between offensive

[17]Glaser (1997) describes the concept as "the key to understanding how in an anarchical international system states with fundamentally compatible goals still end up in competition and war."

and defensive weapons and tactics. As Posen (1993) points out, this was exactly the situation in ex-communist countries that faced the obliteration of their central governments. One might add that this is also the situation that rival ethnic and kin groups face in weak states.

Indeed, in the strategic situation that players face in our model, the two necessary conditions for the security dilemma are present. First, the defensive motive to attack only occurs in an anarchic environment in which players fear they can be preyed on at any time. Hence anarchy is a necessary condition. Second, aggresive and defensive motives are undistinguishable in a very stark way: in the model, the only way of safeguarding the group against attack and potential ethnic cleansing is to mobilize and attack first. As Posen (1993) notes, the tactic of choice in these conflicts tends to be based on small military forces directed against civilians. This technology definitely renders "aggression the most advantageous form of self-defense."

Therefore, despite the fact that the model is not explicitly dynamic, it has the elements that pit groups in a situation of mutual fear tantamount to the onset of a security dilemma. In the model, groups are not reacting to past actions from other groups. However they react to the (rational) beliefs that they might be attacked. As we show in the analysis, the risk dominant criterion predicts that anarchy and self-defense generate the same unfortunate outcome: for $\theta_1 \in \left[\underline{\theta}, \theta^{RD}\right]$, groups "with fundamentally compatible goals still end up in competition and war." Indeed, in this range of $\theta_1$ both groups fight for rationally defensive reasons.

In its traditional exposition, as Jervis and Snyder (1999) describe, the concept of the security dilemma has been shown to suffer from two potential shortcomings:

First, a problem of this concept as a predictive tool is that it tells us what are the conditions necessary for mutual fears to generate conflict but it does not tell us when such conflict will actually happen. Should we expect conflict to result every time anarchy and an offensive advantage coexist? As Fearon and Laitin (1996) point out, in most cases groups are in fact able to live side by side and avoid widespread violence most of the time. This is also true in the anarchic environments of weak states in the developing world. How can we then find out in which situations these mutual fears will actually generate conflict?

Second, as Snyder and Jervis (1999) point out, "in virtually every case [...] the security fears of the parties to civil conflict were intertwined with their predatory goals." This mix of motives constitutes a methodological problem. Most descriptions of the security dilemma point out that this phenomenon exists because one or both groups fear their opponent might be aggressive. In other words, the potential existence of predatory groups is needed for other groups to feel insecure. However, if in fact there are aggresive actors in the system, can we still attribute the existence of conflict to mutual fear?

In our model, these issues find a common answer as agents try to balance the opportunity cost of conflict and the risk of being attacked. Note that in our approach we endogenize the aggressive stance of players: whether a player actually prefers to attack in the absence of a threat does not depend on an unmodeled private "type" but it depends on the current conditions of the economic environment. Players are not aggressive per se. They *become* aggressive when their returns to peace and labor diminish. As we show, the potential existence of these bad economic conditions sets the stage for mutual fears as players try to second guess their opponents. These guesses are complicated by the fact that they are ocurring at different levels simultaneously. A player is not only afraid of the fact that her opponent may want to deviate but she is also concerned that her opponent might react by attacking for defensive motives because she expects a deviation.

As a consequence of this mechanism, we obtain that the security fears that set off conflict are directly related to the economic environment. We also find that the players evolve from *aggresive*, to *security-seekers*, to essentially peaceful as the economic circumstances improve. Hence the framework blends conflicts generated by predatory goals with conflicts generated by security fears in a continuum indexed by the economic situation. For $\theta_1 < \underline{\theta}$ the opportunity cost is so low that both players become aggressive with expansionary goals in mind. For slightly better situations, $\theta_1 \in \left[\underline{\theta}, \theta^{RD}\right]$ the opportunity cost is high enough that peace would be sustainable. However the existing risks derived from mutual fears are too high compared to the opportunity cost and players become embroiled in violence that both players believe is defensive. Finally, for $\theta_1 > \theta^{RD}$, players consider that the return to the economic activity is high enough that they are willing to endure the potential risk of being attacked and hence the security dilemma concerns are not strong enough to generate conflict. Figure 4 illustrates this pattern.

Insert Figure 4 about here

Hence, on the open question of timing we obtain an answer that is supported by the empirical literature: we should expect conflicts to happen when the economic situation is poor. Another advantage of this prediction is that this empirical pattern is in fact much easier to establish than previous emphasis on offense-defense balance.[18] On the methodological question of how to separate conflicts driven by aggressive goals from conflicts driven by security concerns we obtain a framework in which they are really part of the same continuum and security fears set off precisely because there is a contagion effect from states of the economy in which aggression is the only rational activity.

## Conclusion

We have proposed a model where the decision to attack emerges as players try to balance the opportunity cost of violence and the risk of being attacked. This framework can explain the most robust empirical finding in the recent literature on the causes of civil war: bad economic shocks predict conflict. In our model, groups compete today for spoils that will be enjoyed in the future.

[18]See Fearon (1997) and the references thereof.

As a consequence, a bad economic shock reduces the opportunity cost of fighting without reducing the stakes proportionately. Therefore it becomes natural that conflict occurs when the economic situation is poor.

We have also shown that this model transitions from a Prisoner's Dilemma to a Stang Hunt game as economic returns get better. In the latter case, it is well known that equilibrium multiplicity makes prediction a difficult endeavor. At the same time, equilibrium multiplicity allows us to talk about mutual fears generating conflict. However, we standard analysis based on selecting the Pareto efficient equililbrium does not yield the mechanism of mutual fears a described by Jervis (1978). In contrast, we show that those insights can be squared by considering the alternative criterion of risk-dominance, that is specifically designed to take into account players' aversion to strategic risk.

Finally, we have shown that this approach to modelling mutual fears is closely related to the arguments in Posen (1993) and Jervis and Snyder (1999) and can help answer some pending questions in the literature on the security dilemma as applied to civil wars and ethnic conflict. In particular, we show that adding an economic dimension to the anarchic situation helps explain the timing in which mutual fears generate actual conflict. In addition, we also show that the distinction between aggressive and security seeking agents might be a function of prevailing economic circumstances as agents change their attitudes with their opportunity costs and perceptions.

The next step is probably to understand what is the role of the state in a such a situation of miscoordination. One view might be that the state is actually one of the contenders in this fight. A more nuanced and interesting approach would have the state as a third actor, trying to impose peace and solve the coordination problem that pits these groups against each other. From such an approach, we might learn what makes states "weak" in the sense that they cannot avoid the fact that these groups feel they are in an anarchic situation.

## References


Baliga, Sandeep and Tomas Sjöström. 2004. "Arms Races and Negotiations." *Review of Economic Studies* 71(April): 351-369.

Battalio, Raymond, Larry Samuelson, and John Van Huyck. 2001 "Optimization Incentives and Coordination Failure in Laboratory Stag Hunt Games." *Econometrica* 69(May): 749-764.

Carlsson, Hans and Eric Van Damme. 1993. "Global Games and Equilibrium Selection." *Econometrica* 61(September): 989-1018.

Collier, Paul and Anke Hoeffler. 1998. "On Economic Causes of Civil War." *Oxford Economic Papers* 50(October): 563-573.

Collier, Paul and Anke Hoeffler. 2004. "Greed and Grievance in Civil War." *Oxford Economic Papers* 56(October): 563-595.

Collier, Paul and Nicolas Sambanis. 2005a. *Understanding Civil War, Volume 1.* Washington DC: The World Bank.

Collier, Paul and Nicolas Sambanis. 2005b. *Understanding Civil War, Volume 2.* Washington DC: The World Bank.

Cooper, Russell W.,Douglas V. DeJong, Robert Forsythe, and Thomas W. Ross. 1990. "Selection Criteria in Coordination Games: Some Experimental Results." *American Economic Review* 80(March): 218-233.

Dal Bó, Ernesto and Pedro Dal Bó. 2006. "Workers, Warriors and Criminals: Social Conflict in General Equilibrium." *Brown University Typescript.*

De Figueiredo, Rui and Barry Weingast. 1999. "The Rationality of Fear: Political Opportunism and Ethnic Conflict." in Snyder, Jack and Barbara Walter, eds. *Civil War, Insecurity, and Intervention*. New York: Columbia University Press.

Dube, Oeindrila and Juan F. Vargas. 2007. "Commodity Price Shocks and Civil Conflict: Evidence from Colombia." *Harvard University Typescript.*

Ember, Carol R. and Melvin Ember. 1992. "Resource Unpredictability, Mistrust, and War." *Journal of Conflict Resolution* 36(June): 242-262.

Ember, Carol R. and Melvin Ember. 1997. "Violence in the Ethnographic Record: Results of Cross-Cultural Research on War and Aggression." In Martin, Debra L. and David W. Frayer, eds. *Troubled Times: Violence and Warfare in the Past*. Amsterdam: Gordon and Breach.

Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49(Summer): 379-414.

Fearon, James D. 1997. "The Offense-Defense Balance and War since 1648." *Stanford University Typescript*

Fearon, James D. 2005. "Civil War since 1945: Some Facts and a Theory." *Stanford University Typescript.*

Fearon, James D. and David D. Laitin. 1996. "Explaining Interethnic Cooperation." *The American Political Science Review* 90(December): 715-735.

Fearon, James D. and David D. Laitin. 2003. "Ethnicity, Insurgency and Civil War." *The American Political Science Review* 97(February): 75-90.

Flint, Julie and Alex de Waal. 2005. *Darfur: A Short History of a Long War.* London and New York: Zed Books.

Glaser, Charles L. 1997. "The Security Dilemma Revisited." *World Politics* 50(October): 171-201.

Grossman, Herschel I. 1991. "A General Equilibrium Model of Insurrections." *American Economic Review* 81(September): 912-921.

Harsanyi, John C., and Reinhard Selten. 1988. *A General Theory of Equilibrium Selection in Games.* Cambridge, MA: MIT Press.

Herz, John H. 1950. "Idealist Internationalism and the Security Dilemma." *World Politics* 2(January): 157-180.

Hegre, Håvard and Nicholas Sambanis. 2006. "Sensitivity Analysis of Empirical Results on Civil War Onset." *Journal of Conflict Resolution* 50(August): 508-535.

Hidalgo, F. Daniel, Suresh Naidu, Simeon Nichter, and Neal Richardson. 2007. "Occupational Choices: Economic Determinants of Land Invasions." *UC Berkeley Typescript.*

Hirshleifer, Jack. 1995. "Theorizing about Conflict." in Hartley, Keith and Sandler, Todd eds. *Handbook of Defense Economics.* Amsterdam: Elsevier Science.

Hirshleifer, Jack. 2001. *The Dark Side of the Force: Economic Foundations of Conflict Theory.* New York: Cambridge University Press.

Hobbes, Thomas. 2005. *Leviathan, Parts I and II.* Edited by A. P. Martinich. Toronto: Broadview Press.

Jervis, Robert. 1976. *Perception and Misperception in International Politics.* New Jersey: Princeton University Press.

Jervis, Robert. 1978. "Cooperation under the Security Dilemma." *World Politics* 30(January): 167-214.

Jervis, Robert and Jack Snyder. 1999. "Civil War and the Security Dilemma." in Snyder, Jack and Barbara Walter, eds. *Civil War, Insecurity, and Intervention.* New York: Columbia University Press.

Kydd, Andrew H. 1997. "Game Theory and the Spiral Model." *World Politics* 49(April): 371-400.

Kydd, Andrew H. 2005. *Trust and Mistrust in International Relations.* New Jersey: Princeton University Press.

Laitin, David D. 1999. "Somalia: Civil War and International Intervention." in Snyder, Jack and BarbaraWalter, eds. *Civil War, Insecurity, and Intervention.* New York: Columbia University Press.

Miguel, Edward, Shankar Satyanath, and Ernest Sergenti. 2004. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." *Journal of Political Economy* 112(August): 725-753.

Posen, Barry R. 1993. "The Security Dilemma and Ethnic Conflict." *Survival* 35(Spring): 27-47.

Powell, Robert. 1996a. "Bargaining in the Shadow of Power." *Games and Economic Behavior* 15(August): 255-289.

Powell, Robert. 1996b. "Stability and the Distribution of Power." *World Politics* 48(January): 239-267.

Powell, Robert. 1999. *In the Shadow of Power: States and Strategies in International Politics.* Princeton, NJ: Princeton University Press.

Powell, Robert. 2002. "Bargaining Theory and International Conflict." *Annual Review of Political Science* 5(June): 1-30.

Prunier, Gèrard. 1995. *The Rwanda Crisis: History of a Genocide.* New York: Columbia University Press.

Prunier, Gèrard. 2005. *Darfur: The Ambiguous Genocide.* New York: Cornell University Press.

Sambanis, Nicholas. 2002. "A Review of Recent Advances and Future Directions in the Quantitative Literature on Civil War." *Defense and Peace Economics* 13(May): 215-243.

Schelling,Thomas C.1960. *The Strategy of Conflict.* Cambridge, MA: Harvard University Press.

Skaperdas, Stergios. 1992. "Cooperation, Conflict, and Power in the Absence of Property Rights." *American Economic Review* 82(September): 720-739.

Slantchev, Branislav. 2003. "The Principle of Convergence in Wartime Negotiations." *The American Political Science Review* 97(February): 621-632.

Slantchev, Branislav. 2005. "Military Coercion in Interstate Crises." *The American Political Science Review* 99(November): 533-547.

Snyder, Jack. 2002. "Anarchy and Culture: Insights from the Anthropology of War." *International Organization* 56(Winter): 7-45.

# Appendix A

**Proof of Lemma 1.** Assume that $(a, p)$ is an equilibrium for some realization $\hat{\theta}_1$. For a player to find it optimal to attack when the other one is peaceful, it has to be that

$$\delta\pi\hat{\theta}_1 + (1+\delta)T > \pi\hat{\theta}_1 + T$$

which can be rewritten as

$$\hat{\theta}_1 < \frac{\delta T}{(1-\delta)\,\pi}$$

At the same time, for a player to passively accept an attack, it must be true that

$$(1-\delta)\,\pi\hat{\theta}_1 + (1-\delta)\,T - S > T - W$$

which can be rewritten as

$$\hat{\theta}_1 > \frac{S - W + \delta T}{(1-\delta)\,\pi}$$

For $\hat{\theta}_1$ to exists, it has to be that

$$\frac{\delta T}{(1-\delta)\,\pi} > \frac{S - W + \delta T}{(1-\delta)\,\pi}$$

which can only be true if $S - W < 0$. Hence no such equilibrium can exist under Assumption 1.■

**Proof of Proposition 1.** Let us check the best response functions. For a player that expects to be attacked, her best reply is to attack as long as

$$(1-\delta)\,\pi\theta_1 + (1-\delta)\,T - S < T - W$$

or, by rearranging

$$\theta_1 < \frac{S - W + \delta T}{(1-\delta)\,\pi} = \bar{\theta}$$

Since the realization is common to both players, $(a, a)$ is an Nash Equilibrium if and only if $\theta_1 < \bar{\theta}$. Similarly, for a player not to attack a peaceful opponent, it has to be that

$$\delta\pi\theta_1 + (1+\delta)T < \pi\theta_1 + T$$

and again, by rearranging

$$\theta_1 > \frac{\delta T}{(1-\delta)\,\pi} = \underline{\theta}$$

Therefore, $(p,p)$ is a Nash Equilibrium of the game if and only if $\theta_1 > \underline{\theta}$. Moreover, it is easy to show that

$$\underline{\theta} = \frac{\delta T}{(1-\delta)\,\pi} < \frac{S - W + \delta T}{(1-\delta)\,\pi} = \bar{\theta}$$

where the inequality is immediate from Assumption 1. Since Lemma 1 precludes the existence of assymetric equilibria, the structure of Nash Equilibria in the game has to follow Proposition 1.■

**Proof of Proposition 3.** Substitute $\underline{\theta}$ into (3) to obtain

$$\theta^{RD} = \frac{\delta}{1-\delta}\beta\Theta + \frac{S-W}{2(1-\delta)\pi}$$

From this expression, together with Assumption 1, it is clear that $\theta^{RD}$ is increasing in $\delta$ and $S$, and it is decreasing in $\pi$ and $W$.■

**Proof of Proposition 4.** Take the following modified version of the game. Assume that player $A$ plays rows and player $B$ plays columns. The following matrix describes the payoffs for $A$,

| | $p$ | $a$ |
|---|---|---|
| $p$ | $\pi_A\theta_1 + \beta\pi_A\Theta$ | $(1-\delta)\,\pi_A\theta_1 + (1-\delta)\,\beta\pi_A\Theta - S$ |
| $a$ | $\delta\pi_B\theta_1 + \beta\pi_A\Theta + \delta\beta\pi_B\Theta$ | $\beta\pi_A\Theta - W$ |

while this matrix contains the payoffs for $B$.

| | $p$ | $a$ |
|---|---|---|
| $p$ | $\pi_B\theta_1 + \beta\pi_B\Theta$ | $\delta\pi_A\theta_1 + \beta\pi_B\Theta + \delta\beta\pi_A\Theta$ |
| $a$ | $(1-\delta)\,\pi_B\theta_1 + (1-\delta)\,\pi_B\Theta - S$ | $\beta\pi_B\Theta - W$ |

Note that when a group mobilizes and its opponent plays peace, it seizes an amount of land $\delta$ that has the productivity of the other group. Besides these differences in the payoffs, the game is identical to the one described in section II. We can proceed straight to the identification of $\theta^{RD}$.

The product of deviations from the $(p, p)$ equilibrium equals

$$\left(\pi_A\theta_1 + \beta\pi_A\Theta - \left[\delta\pi_B\theta_1 + \beta\pi_A\Theta + \delta\beta\pi_B\Theta\right]\right)\left(\pi_B\theta_1 + \beta\pi_B\Theta - \left[\delta\pi_A\theta_1 + \beta\pi_B\Theta + \delta\beta\pi_A\Theta\right]\right)$$

In the same manner, we can express the product of deviations from the $(a, a)$ equilibrium

$$\left(\beta\pi_A\Theta - W - \left[(1-\delta)\,\pi_A\theta_1 + (1-\delta)\,\beta\pi_A\Theta - S\right]\right)\left(\beta\pi_B\Theta - W - \left[(1-\delta)\,\pi_B\theta_1 + (1-\delta)\,\pi_B\Theta - S\right]\right)$$

For small $\Delta = \pi_A - \pi_B$, the first expression is increasing in $\theta_1$ and the second expression is decreasing. By equalizing these two expressions we can obtain an implicit determination of $\theta^{RD}$.

$$-\delta\Delta^2\left(\left[\theta^{RD}\right]^2 + \theta^{RD}\beta\Theta\right) + \theta^{RD}\left(1-\delta\right)\left(\pi_A + \pi_B\right)\left(S - W\right) = \left(S - W\right)^2 + \delta\beta\Theta\left(\pi_A + \pi_B\right)\left(S - W\right)$$

Note that if $\Delta = 0$ this expression yields exactly $\theta^{RD}$ from (3). Totally differentiating one obtains:

$$\frac{\partial\theta^{RD}}{\partial\Delta} = \frac{2\delta\left(\left[\theta^{RD}\right]^2 + \theta^{RD}\beta\Theta\right)}{\frac{1}{\Delta}\left(1-\delta\right)\left(\pi_A + \pi_B\right)\left(S - W\right) - \delta\Delta\left(2\theta^{RD} + \beta\Theta\right)}$$

Which is strictly positive for $\Delta \in (0, \xi)$, for small enough $\xi > 0$.■

**Proof of Proposition 5.** See Carlsson and van Damme (1993).■

# Appendix B

Suppose that two players, $A$ and $B$, are considering the civil war payoff matrix (1). Suppose that $A$ believes that $B$ will attack with probability $\alpha \in [0, 1]$. Given $\alpha$, $A$ will decide to play $p$ if

$$\left(1-\alpha\right)\left(\pi\theta_1 + T\right) + \alpha\left(\left(1-\delta\right)\pi\theta_1 + \left(1-\delta\right)T - S\right) > \left(1-\alpha\right)\left(\delta\pi\theta_1 + (1+\delta)T\right) + \alpha\left(T - W\right)$$

or, with some manipulation

$$\frac{1-\alpha}{\alpha}\frac{\left(1-\delta\right)\pi\theta_1 - \delta T}{S + \delta T - W - \left(1-\delta\right)\pi\theta_1} > 1 \tag{4}$$

This expression can be interpreted in an intuitive way. The first term $\frac{1-\alpha}{\alpha}$ is the relative probability of actually playing the right action when playing $p$ versus playing war $a$. Playing $p$

is the right action when the opponent is not attacking, and this happens with probability $1-\alpha$. Playing $a$ is the right action when the opponent attacks, with probability $\alpha$. Hence the ratio is the relative probability of *not* making a mistake when playing $p$.

The second term captures the relative benefits of playing the right action. The numerator displays the gains from staying in the $(p,p)$ equilibrium. Note that when the opponent plays $p$, $A$ obtains $\pi\theta_1 + T$ by responding with $p$ and $\delta\pi\theta_1 + (1+\delta)T$ by attacking. The difference between these two payoffs, which has to be positive if $(p,p)$ is a Nash Equilibrium, is the relative benefit of playing $p$ in response to $p$. In the same way, the denominator features the benefit from staying in the $(a,a)$ equilibrium.

Hence, the expression has a simple interpretation: when the relative probability of getting it right by playing $p$ times the relative gains of getting it right by playing $p$ is greater than one, $A$ should play $p$. To apply this criterion, $A$ only an assessment of $\alpha$.

To obtain this assessment, $A$ tries to look at the game from $B$'s point of view. What does $A$ imagine $B$ thinking that $A$ will do? Applying the principle of insuficient reason, Harsany and Selten assume that player $A$ has uniformly distributed beliefs on how $B$ sees her probability of attacking. In other words, $A$ thinks that $B$ thinks that $A$ will attack with probability $\beta\tilde{}U\left[0,1\right]$. Under this assumption, $A$ knows that $B$ will attack if $\beta > \hat{\beta}$ where $\hat{\beta} = \frac{(1-\delta)\pi\theta_1 - \delta T}{-2\delta T + S - W}$.[19]

Given this and the uniformity assumption, $A$'s assessment of the probability that $B$ will attack equals $Prob(\beta > \hat{\beta}) = 1 - \hat{\beta} = \frac{-\delta T + S - W - (1-\delta)\pi\theta_1}{-2\delta T + S - W}$. By substituting this in expression (4) in the place of $\alpha$, one obtains the risk dominance criterion established by Harsanyi and Selten (1988):

$$\left(\left(1-\delta\right)\pi\theta_1 - \delta T\right)\left(\left(1-\delta\right)\pi\theta_1 - \delta T\right) > \left(S - \delta T - W - \left(1-\delta\right)\pi\theta_1\right)\left(S - \delta T - W - \left(1-\delta\right)\pi\theta_1\right)$$

In words, $(p,p)$ risk dominates if it is associated with the largest product of deviation losses.

[19]This threshold is calculated from

$$\left(1-\hat{\beta}\right)\left(\pi\theta_1 + T\right) + \hat{\beta}\left(\left(1-\delta\right)\pi\theta_1 + \left(1-\delta\right)T - S\right) = \left(1-\hat{\beta}\right)\left(\delta\pi\theta_1 + (1+\delta)T\right) + \hat{\beta}\left(T - W\right)$$

## Appendix C

Note that as $\epsilon$ goes to 0, both players know almost exactly what the real state of the economy, $\theta_1$, is. However, not knowing it exactly makes a big difference. Since there is no longer direct knowledge of $\theta_1$, a strategy is now a mapping from $s_i$ to $p$ or $a$. In other words, it is a complete contingent plan that assigns an action to every possible signal that a player might receive.

For every $s_i$ player $i$ updates her beliefs on $\theta_1$ and on $s_j$. Her posterior beliefs on $\theta_1$ are $U\left[s_i - \frac{1}{2}\epsilon, s_i + \frac{1}{2}\epsilon\right]$. And the conditional distribution of $s_j$ is

$$\begin{aligned} P(s_j &> \bar{x}/s_i) = \frac{1}{2} + \frac{(s_i - \bar{x})^2}{2\epsilon^2} + \frac{s_i - \bar{x}}{\epsilon} \qquad \text{if } \ s_i \leq \bar{x} \\ P(s_j &> \bar{x}/s_i) = \frac{1}{2} - \frac{(s_i - \bar{x})^2}{2\epsilon^2} + \frac{s_i - \bar{x}}{\epsilon} \qquad \text{if } \ s_i \geq \bar{x} \end{aligned}$$

This density is symmetric around $s_i$ with support $[s_i - \epsilon, s_i + \epsilon]$. For every $s_i$ when $\epsilon$ is small, player $i$ knows that the signal that her opponent has received is very close to hers, but it can still be slightly above or slightly below $s_i$. This strategic risk does not disappear as $\epsilon \to 0$.

Note that any equilibrium needs a minimum of a switching point between playing $a$ and playing $p$. To see this, note that for $s_i < \underline{\theta} - \frac{1}{2}\epsilon$, player $i$ plays $a$ for sure as she knows she is in the area where opportunity costs are so low that violence is a best reply no matter what $j$ does. For $\epsilon$ small enough, the fact that $Pr(\theta_1 < \underline{\theta}) > 0$ ensures that such signals exist. Likewise, for $s_i > \bar{\theta} + \frac{1}{2}\epsilon$, player $i$ will play peace. It is clear then that at some point the player has to switch from attacking to playing peace. However, for signals close enough to a potential switch point, the players are not sure anymore whether their signal is at the same side of the threshold as the signal of the opponent. This means that players have to cope with some positive probability of being attacked and not being prepared. A switch point, therefore can only occur when a player is indifferent between $p$ and $a$ given the strategy of the opponent.

Imagine that $\bar{x} \in \left(\underline{\theta}, \bar{\theta}\right)$ is a threshold at which $j$ switches from $a$ to $p$. What is $i$'s best response?

Given that $\bar{x}$ is in the central area, $i$ will try to mimic the actions of $j$ around the threshold. As a function of $s_i$, her return to playing $p$ is:

$$P(s_j > \bar{x}/s_i)\left(\pi E(\theta_1/s_j > \bar{x}, s_i) + T\right) + \left(1 - P(s_j > \bar{x}/s_i)\right)\left((1-\delta)\,\pi E(\theta_1/s_j < \bar{x}, s_i) + (1-\delta)\,T - S\right)$$

Also, her return to playing $a$ is:

$$P(s_j > \bar{x}/s_i)\left(\delta \pi E(\theta_1/s_j > \bar{x}, s_i) + (1+\delta)T\right) + \left(1 - P(s_j > \bar{x}/s_i)\right)(T - W)$$

The first expression is clearly increasing in $s_i$, while the second one is decreasing. Hence there is a threshold $\bar{s}$ at which player $i$ will switch. By equating the previous expressions and simplifying, one obtains:

$$P(s_j > \bar{x}/\bar{s})(S - W) + \pi(1-\delta)\bar{s} = S - W + \delta T \tag{5}$$

This equation determines the optimal switching point of $i$ as a function of the switching point of $j$. But obviously, for these switching points to constitute an equilibrium, $\bar{x}$ needs to be a best response to $\bar{s}$. The condition for this is:

$$P(s_j > \bar{s}/\bar{x})(S - W) + \pi(1-\delta)\bar{x} = S - W + \delta T \tag{6}$$

It is easy to show that these two conditions imply $\bar{x} = \bar{s}$. Suppose not. Without loss of generality, assume that $\bar{x} > \bar{s}$. This means that $P(s_j > \bar{s}/\bar{x}) > \frac{1}{2}$ and $P(s_j > \bar{x}/\bar{s}) < \frac{1}{2}$. But then the left hand side of (6) necessarily has to be greater than the left hand side of (5) which is not possible as both equal the same expression. $\bar{x} = \bar{s}$ implies $P(s_j > \bar{x}/\bar{s}) = \frac{1}{2}$ and hence one immediately obtains from any of these expressions $\bar{x} = \bar{s} = \theta^{RD}$.

While this is only a sketch of the full proof, it shows that the fact that players are not certain they have signals at the same side of a threshold, reduces the set of possible thresholds to a single one: that in which the opportunity cost of fighting is just high enough that players are indifferent between attacking for security motives and staying peaceful, given a $\frac{1}{2}$ chance of being attacked.