Onset of Armed Conflict: A New List for the period 1946–2004, with Applications

Håvard Strand[a]

hs@prio.no

Department of Political Science, University of Oslo &
the Centre for the Study of Civil Wars, PRIO

Abstract: This article introduces a new list of onsets based on the Uppsala/PRIO Armed Conflict Dataset (ACD) and is constructed with the ambiguous nature of conflict definitions in mind. The new dataset allows for extensive robustness checks partially addressing the problem presented and discussed by Sambanis (2004). These capacities are demonstrated through a re-analysis of Hegre et al. (2001), which concludes that the finding of the effect of proximity to regime change on risk of conflict is robust, while the effect of political regimes is dependent on the definition of onset used.

[a]I am grateful to the Centre for the Study of African Economies, Oxford University for hosting me during the work with this paper, and to Jim Fearon, Scott Gates, Nils Petter Gleditsch, Håvard Hegre, Anke Hoeffler, Alexander Moradi, Henrik Urdal and Gudrun Østby for comments and discussions. Special thanks go to Stine Thomassen and Christin Ormhaug for excellent and enduring research assistance.

Baldrick: No, the thing is: The way I see it, these days there's a war on, right? and, ages ago, there wasn't a war on, right? So, there must have been a moment when there not being a war on went away, right? and there being a war on came along. So, what I want to know is: How did we get from the one case of affairs to the other case of affairs?

Edmund: Do you mean "How did the war start?"

Baldrick: Yeah.

"Blackadder Goes Forth" (Curtis et al., 1998: 414f)

Introduction

It is a well known problem in the conflict literature that statistical results are not always robust. This lack of consistency fuels academic quarrels and causes confusion among policy makers (Mack, 2002). Hegre & Sambanis (2006) draw our attention to the operationalization of the explanatory variables. They show that seemingly arbitrary choices of independent variables often produce significantly different conclusions. Furthermore, Sambanis (2004) shows that alternating between different conflict datasets can produce very different results. The conflict onset literature should build upon these insights, but so far there has not been a framework enabling such advance. This article focuses on the dependent variable, presenting a new list of conflict onsets that enables each individual user to easily test the robustness of his/her findings. Through an analysis of political regimes and instability the article illustrates how this can be done and why it should be required for future work in this area.

The current empirical literature on the causes of armed conflicts is to a large extent based on predefined conflict datasets (e.g. Sambanis, 2004; Fearon & Laitin, 2003; Gleditsch, 2004), which by nature rely on a number of arbitrary coding decisions (Sambanis, 2004). Rather than fixing these decisions once and for all, the dataset presented here opens up two elements of the definition to the individual user's discretion: the number of fatalities marking the threshold of inclusion, and the minimum period of peace required before we code a new conflict between old adversaries. Furthermore, this dataset builds on the Uppsala/PRIO Armed Conflict Dataset (Gleditsch et al., 2002) extending the

definition of conflict in order to produce an onset list – a job hitherto left to each researcher.

The article first discusses the conceptual aspects of conflict definitions that must be more or less arbitrarily defined, and then presents a new definition of conflict onset. The aspects of the definition that can be altered by the users are given particular focus and the framework for a robust analysis of onset is presented. The final section presents a re-investigation of the seminal contribution by Hegre et al. (2001) on the relationship between political instability and conflict onset. They showed that both semi-democratic regimes and proximity to political instability are associated with increased risk of conflict. Through alternating between different definitions of conflict onset I show that the effect of political instability remains robust while the war-proneness of semi-democracies is more dubious.

Conceptualizing Conflict Onset

The conflict data available today can perhaps be divided into two. One group relates closely to the Correlates of War (COW) project (Small & Singer, 1982; Sarkees, 2000) and focuses on 'war', whereas the other group is more inclusive and includes 'armed conflict'. In the conflict literature, the term 'war' has been associated with a conflict with more than 1000 fatalities. The term 'armed conflict' is more loosely defined, but seldom includes conflicts with less than 25 fatalities.

Among the COW derivates, Sambanis (2004); Collier & Hoeffler (2004); and Gleditsch (2004) are perhaps the most prominent. The alternative camp includes among others the Kosimo project (Pfetsch & Rohloff, 2000): Marshall (2005); and the Uppsala Conflict Data Program (Harbom, 2005; Harbom, Högbladh & Wallensteen, 2006; Gleditsch et al., 2002). The latter includes the Uppsala/PRIO collaboration, which this study builds on. Also building on Uppsala/PRIO are Cunningham, Gleditsch & Salehyan (2005) and Gates & Strand (2006). Fearon & Laitin's (2003) dataset is somewhere in between, as it focuses on wars but applies criteria different from COW.

Common for all these projects is that conflict is seen as distinct from other forms of violence, such as crime, one-sided violence (genocides and politicides), and communal violence (i.e. fighting among different non-state actors). This distinction is expressed both through the requirement of a political incompatibility (as a pure economic 'incompatibility' would be seen as criminal activity) and the requirement of an organized opposition. The internal ranking of these two requirements exemplifies the problem highlighted by Sambanis (2004). It is unclear whether incompatibility or organization should have priority when coding conflicts. Most datasets give priority to the incompatibility, as the alternative would imply a new onset every time a new organization is formed. Cunningham, Gleditsch & Salehyan (2005) is perhaps the best example of a dataset prioritizing the organizational aspect, but their analysis focuses on conflict duration rather than onset.

The Uppsala definition of armed conflict is a contested incompatibility, involving at least two organized parties of which at least one is a recognized government, over a stated political incompatibility, where at least 25 people are killed in battle-related circumstances (Strand et al., 2005: 4). The Uppsala definition gives clear priority to the incompatibility, to the extent that all fighting over the same incompatibility is considered one conflict, unless there is a 'complete change on the opposition side'. This criterion has proven difficult to incorporate.

Gates & Strand (2006) introduced additional criteria to better separate between different conflicts over time. By these criteria, a conflict-year including previously active parties constitutes a new conflict if it is preceded by at least two consecutive years of inactivity. Any onset of violence between a government and a new opposition organization is seen as a new conflict if the conflict outbreak occurs at a time when no other organizations are active within that specific incompatibility. Gates & Strand (2006) give substantially higher priority to the organizational dimension of conflict. The Gates & Strand dataset was developed to analyze duration, but is also relevant for the definition of onset.

Sambanis' (2004) discussion of conceptual differences among conflict datasets highlights three dimensions: The casualty threshold applied; the

definition of start and end dates; and the distinction between interstate, intrastate, and extrastate conflict. Sambanis shows that these three dimensions all have coding criteria that are only loosely connected to conflict theory, and therefore involve a substantial element of arbitrariness. This article focuses on thresholds and dates leaving the nominal categorization to others.

Defining the necessary number of casualties is clearly difficult. There is a qualitative difference between violent and non-violent political conflicts, but there is no such difference between 24 and 25 fatalities per year. Nor is there necessarily a qualitative difference between 25 and 500 fatalities, although the quantitative difference is quite large. Furthermore, once a conflict has been included, we must decide when it started and when it stopped. We might want to date the start of the conflict to the point in time where the first casualty was observed, or to the event that lead to the fulfillment of all the criteria. There are good arguments for both these alternatives, but no argument has so far been accepted as conclusive. The arguments for each option are discussed in detail in the next section.

The conceptual variation shown by Sambanis (2004) is a likely cause for the great variations of the findings in the literature. Sambanis (2004) analyzes this variation by regressing the same set of Right Hand Side (RHS) variables on different conflict datasets. Hegre & Sambanis (2006) analyzes the variation by running different sets of RHS variables on two different conflict dataset. Both these approaches are interesting and revealing, but they are computationally intensive and very time-consuming. The alternative proposed here both takes into account the variation caused by conceptual ambiguity and is easily adoptable at the same time.

Finally, there is a need for precise dating of conflict onsets. A conflict onset is something which happens, and can therefore be dated. The causes of such events can often also be seen as events, for example as an abrupt political change in the form of a coup d'état. In such circumstances the timing of the events are of crucial importance. The complementary category, states, captures more stable aspects, such as a society being in a state of war or a state of poverty. While the standard country-year format is well suited for analyzing the

effect of states on the likelihood of conflict, it is not capable of distinguishing between cause and consequence if two events happen within one single year. Focusing on persistent variables such as poverty or natural resources inform us about what countries are most likely to see conflict, but due to the stable nature of these variables, they can tell us little about the exact timing of conflicts. Therefore we need to focus more on conflict onsets as events, and date them as precise as possible.

A New Onset Definition

The list of onsets defined here builds on the Uppsala/PRIO Armed Conflict Dataset (ACD) (Gleditsch et al. 2002; Harbom, Högbladh & Wallensteen, 2006). ACD defines conflict as a violently contested incompatibility between two or more parties of which one is a recognized government, and where 25 or more persons are killed in battle-related incidences within a calendar year. The definition separates between annual observations where more than 1000 persons are killed – labeled 'war' – and those where the fatality estimate was between 25 and 1000 – labeled 'minor'. A third label, 'intermediate', is used where the cumulative number of fatalities is above 1000, while the annual count for that specific year is below 1000. The list presented here includes information on whether the conflict onset results in a 'war', an 'intermediate' conflict, or a 'minor' conflict. The casualty threshold will be referred to as δ and is user-adjustable.

Gates & Strand (2006) added two aspects to the Uppsala/PRIO definition that are relevant for this study. The first addition is mapping the event which should be counted as the initiation of conflict. Gates & Strand define two different start dates: the first is the initial fatality in the series of fatalities that eventually met the threshold of 25 battle-related deaths (BRD), and the second is the date when this threshold was breached.

Gates & Strand's second addition is the question of intermittent conflict. Many conflicts show intermittent patterns of violence, in which violent periods can be separated by long periods of little or no fighting. It is difficult to theoretically define when an intermittent period is long enough to warrant a

new onset, and it relies on what Sambanis (2004) refers to as ad hoc definitions. Rather than cementing one ad hoc definition this dataset allows the user to easily vary this definition of inclusion between two and ten years. This intermittency threshold will throughout this article be referred to as ε.

Building on the conflict definition presented in Gleditsch et al. (2002) and in Gates & Strand (2006), I propose the following operational definition of conflict onset:

§1. The onset of an internal conflict is defined at the first day where the Uppsala criteria were breached. In most cases this is the day when the 25th person was killed.
    a. If the empirical foundation is ambiguous or insufficient, the day of onset is coded as early as possible. I.e. if the sources only inform us of the month or the year, the first day of that month or year is coded.
    b. If a new opposition organization joins an active conflict over a specified incompatibility, this is not recorded as a new onset.
    c. Regardless of the value of δ chosen, the date recorded will be according to §1.

§2. A conflict between a government and an opposition organization is to be coded according to §1 if all previous conflict over the same incompatibility has been settled, and the opposition organization in question is unrelated to any other organizations active the last ten years.

§3. Two conflict-years over the same incompatibility, with the same active organizations should be treated as separate conflicts if they are separated by an intermittent period of ε years, where ε is between two and nine years.
    a. The intermittency threshold ε can be increased up until 9 years by the users of the dataset.

This definition chooses the day when the Uppsala criteria are reached as the crucial date in the process from peace to conflict. The only alternative that would make sense is the day of the first use of fatal violence. This date would

serve as a guarantee against a spurious relationship, e.g. if the first fatality caused a chain of events that in turn caused the 25th fatality. However, there are both substantial and pragmatic reasons supporting the latter date. Quite often, both the first and the 25th fatality occur on the same date. If this is not the case, it is much more challenging to positively identify the initial fatality, and hence the measurement error would be significantly larger if we chose the initial fatality. On a more substantive note, the two events can be separated by a substantial amount of time, in some cases years. Choosing the first fatality would create the need for another arbitrary clarification: When is the first fatality a single, isolated event and when is it connected with a successive escalation?

When it comes to §1c, pragmatism must be revoked again. Restricting the casualty threshold to conflict with at least one year with more than 1000 BRD raises the same question. It might seem as if choosing the date of the 1000th fatality would be the logic conclusion from the discussion above. However, this would introduce a substantial element of endogeneity into the analyses. By the time of the 1000th fatality, a conflict is very likely to have had negative consequences, and this negative impact will then be reflected in the regressors explaining conflict. Thus, one should always code conflict onset at an early stage, when the explanatory variables are still unaffected.

The dataset presented in Gleditsch et al. (2002) frequently views several different violent periods involving different opposition organizations as the same conflict, while the definition of Gates & Strand (2006) is more likely to view these periods as different conflicts. An example is the events of 1958–1959 in Iraq. Gleditsch et al. views the coup in 1958 and the attempted coup in 1959 as involving the same organizations, while Gates & Strand holds the nationalist, conservative group of officers attempting a coup against the Qasim government as distinct from the latter's successful revolution the year before. Such difference in interpretation is the main explanation of why Gates & Strand counts more conflicts than those reported in Gleditsch et al. and later updates. Interpreting the 'Complete change on the opposition side' criterion can be difficult. Consider two parties fighting over the government of a country, where

the opposition organization is capable of overtaking the capital, and thereby the government, but not eradicate its adversary. Both in the Republic of Congo and Afghanistan this situation is repeatedly relevant. A strict and dogmatic interpretation of the Uppsala criteria could lead a researcher to think that all complete changes on the opposition side constitute a new conflict, and a new onset. When the alliance of Afghan warlords and a US-lead coalition won control over Kabul from the Taliban, this should be coded as a new onset according to this logic, even though the war went on. This would not be a valid operationalization of what we theorize as a conflict onset.

A more flexible operationalization is needed. Gates & Strand (2006) solve this problem requiring the previously active conflict dyad(s) to have settled their incompatibility before any new conflict can be coded within that incompatibility. It follows that the new opposition organization, whose emergence caused the coding of a new conflict, should not be organizationally related to any previous active party.

The terms 'settled' and 'previously active' are somewhat ambiguous, and must also be defined. Gates & Strand (2006) defines settlement as either a lasting solution to the conflict, either military or peaceful, or alternatively a period of ten years of no activity. Previously active is therefore also here defined as listed by Uppsala as active in the preceding ten-year period.

The distinction between new and reoccurring conflict is important because the emergence of a totally new contender over a given incompatibility can happen as soon as any previously active conflict has been settled. This is in contrast to a situation in which the new contender is seen as a continuation of an old party, for instance in the form of a splinter group, where the situation would be interpreted using the rules of intermittent conflict.

However, if one also wants to open up for the possibility of reoccurring conflict, one needs §3. How long should a conflict be inactive before one allows it to restart, rather than interpreting the episode of violence in question as the continuation of the previous conflict? There is no authoritative answer to this, and the choice should as far as possible be left to the individual user's discretion. It is difficult to code reliable data below a two-year limit of inactivity. The upper

boundary is effectively defined by §2 at ten years, since any period of inactivity longer than ten years will produce a new onset regardless of organizational matters.

Ambiguous Definitions and Robust Analysis

The definition of onset presented in the previous section is more inclusive than previous alternatives, as it includes more conflicts due to lower casualty threshold and more onsets due to a short intermittency threshold of two years. This list includes 275 entries, whereas Fearon & Laitin (2003: 75) and Sambanis (2004: 831) have 127 and 145, respectively. Since it is easier to remove existing entries than add new ones, an inclusive list has an advantage regarding robustness analysis.

This author seconds and applauds Sambanis' (2004) insights regarding ad-hoc definitions of important coding criteria. We have not yet sufficient theoretical basis for deriving all aspects of a complete conflict definition, and this shortcoming should inspire future research. Hence, we should question the neutrality of the ad-hoc definitions. Of the three dimensions discussed by Sambanis, this dataset opens the definition of the casualty threshold and intermittency threshold to each user.

The alteration of the threshold is limited to intensity as defined by the Uppsala Conflict Data Project (Harbom & Wallensteen, 2005). Incidence of conflict within a calendar year is given one of three intensity codes.[1] This variable is not ordinal, but consists in fact of two dimensions. The first dimension is whether the conflict caused more than 1000 BRD in that calendar year or between 25 and 1000 BRD. If the conflict failed to produce 25 BRD, it will of course not be included for that particular year. Furthermore, there is a cumulative category that separates between conflicts that has produces more than 1000 BRD or not over the entire span of its duration. It follows from this definition that all observations where the 1000 BRD limit is met will also meet the cumulative 1000 BRD threshold. The distinction between the annual and

[1] This has been changed somewhat in the 2006 release of these data, but the information is still contained, albeit in two different variables.

cumulative dimensions is therefore only relevant if the annual threshold is below 1000 BRD. The three categories can be summed up in Table 1:

Table 1: Three Intensity Categories

	25–999 pr. year	**1000+ in at least one year**
25–999 in total	Minor	*impossible*
1000+ in total	Intermediate	War

This information can be used to select subsets of onsets. Based on the maximum intensity coding of a conflict over its entire duration, we can classify conflicts into three categories:

1. Conflicts that never reached the cumulative 1000 BRD threshold.
2. Conflicts that reached the cumulative 1000 BRD threshold, but never reached the annual 1000 BRD threshold.
3. Conflicts that reached the annual 1000 BRD threshold at least for one year and therefore by definition also reached the cumulative threshold.

By selecting either all onsets (category 1+2+3), by the cumulative threshold (category 2+3) or only the major wars (category 3), we can get an impression of what the effect of the threshold criterion has on our conclusions.

Intermittent conflicts present unique problems to definitions of conflict onsets. If a given conflict is inactive for a significant period of time, should the observation of renewed violence between the previously engaged parties be coded as an onset? This depends on the duration of the inactive period and, as with the casualty threshold, an increase of one unit will never make a difference qualitatively.

If we shorten the intermittency period, we can increase our sample of wars, but we might also erroneously include onsets that in reality are ongoing conflicts. This can happen either as a function of the warring parties' prolonged retirement from battle in order to gain strength, or can be a result of incomplete reporting. Since the Uppsala Conflict Data Program has a very strict criterion regarding what information is accepted as credible, information shortage is likely to result in wrongful exclusion of observations. By increasing from two to ten the number of years of inactivity needed for a conflict to be coded as ended, we can test the robustness of our results against this ad hoc parameter.

An analysis will be robust if either the results are unaffected by changing the casualty and intermittency threshold, or if the changes caused by changing the definition can be explained and derived theoretically. In order to verify such robustness we need a simple framework for adjusting the thresholds. This framework must not only focus on which onsets to include at the various thresholds, but also help us define the control group.

Let w be the set of all onsets defined at the most inclusive level, and let us denote all subsets of w as $w^{\delta,\varepsilon}$, with δ as the intensity threshold and ε as the minimum intermittency period. The most inclusive set of onsets is 25 BRD and a two-year intermittency criterion, which will be denoted $w^{1,2}$. Onset analysis is basically about comparing conflict onsets with a control group, comprised of units that could have experienced an onset, but which for various reasons did not. How should we observe this control group and how does it relate to w? One possibility is to define a complimentary set p as a set of all observations that did not experience an onset. The combination $N=w \cup p$ would then be a complete set of all observations. If we define an observation to be a country-year, we have the standard data structure used in most quantitative studies of conflict.

Despite the popularity of the country-year approach, there is no theoretical justification for this specific design. We could use monthly periods of five-year periods instead, and it would no more or less correct. It is also problematic that when we increase either δ or ε, we are in fact moving observations from w to p. Thus, changing the set of onsets implies changing the composition of the control group as well.

A robustness analysis for a country-year design would consist of creating several variables with to substitute the dependent variable. A result is robust if we get the same results using any value for δ and ε. But if covariates of $w^{1,2}$ are similar to those of $w^{3,9}$ then this logic fails. The dependency between the set of onsets and the control group can affect the covariates of the control group to the extent that the results might change, not because of differences within w but because the proposed robustness check would move elements of w over to p. This shortcoming warrants an alternative data structure.

When we theorize about conflict onset, we often treat it as if it happens suddenly. This is especially true for event-based arguments, such as that of Hegre et al. (2001), where regime change is seen as a direct cause of conflict. This necessitates that the regime change happened before the conflict onset, even if there are only a matter of days between two events. In a country-year design, it is hardly possible to consistently code events that happen within a country-year unit as anything but simultaneous events.

Hegre et al. (2001), drawing on Raknerud & Hegre (1997), develops a model with the control group defined as all countries which did not experience a conflict onset at the exact day when one country actually did experience one. This can be thought of as a country-day design, but the control group is not all days when no event took place. Instead the control group is the set of all countries where there was no conflict onset, observed at the day where there in fact was an onset in one country. Thus, the dataset consists of w onsets, with a set of observations equal to $N = \bigcup_{i=1}^{w} z_i$, where z_i is the set of independent countries at each point in time i. This design is easy to justify, as the relationship between N and w is defined through the international system of independent states.

In a country-year design, a robustness check consisting of eliminating a given onset would increase the size of the control group by one unit and similarly reduce the onset group. By contrast, the Raknerud-Hegre design would remove not only the onset in question from the sample, but all elements of the control group sampled at the same point in time. Thus, no observation is moved from one group to the other. Thus, if the set of onsets w is a homogeneous set, an analysis of a subset of w should yield the same results as an analysis of w.

A more practical justification for the Raknerud-Hegre design is that any robustness check can be applied as a filter in a regression analysis, which effectively filters both the onset and control groups at the same time. The analysis presented in this paper exemplifies how this can be done.

Political Regime, Political Stability, and Conflict

Political instability and semi-democratic regimes are often associated with onset of armed conflicts and civil wars.[2] Hegre et al. (2001) discusses whether the inverted U-curve relationship reported between political regime and conflict onset is due to the fact that semi-democracies are less stable than other regimes, but find that this is not the case. Gates et al. (2006) finds that semi-democracies are indeed less stable than both autocracies and democracies, while Strand (2006) argues that Hegre et al.'s finding of an inverted U-curve is due to an endogeneity problem in the Polity dataset.

Theories regarding regime change in general and democratization in particular argues that engineering political systems is a risky business. Democratization in particular tends to be partially successful, often resulting in a semi-democratic regime. Yet the causal link is still somewhat unclear, and therefore what the link between regime change and conflict initiation consist of. Any outbreak of armed conflict is likely to be considered a regime change. A further consequence of this relationship is that the end of an armed conflict in itself will constitute a regime change, often in the direction of a semi-democratic regime. The literature on the causes of armed conflict has consistently shown that post-conflict periods are more at risk of experiencing renewed conflict than are similar situation without any preceding conflict (Collier et al. 2003). Is the relationship between proximity to regime change and onset of conflict merely a part of the conflict trap?

The shorter the time between an onset and the end of the previous conflict, the more likely is the cause part of the conflict trap. The robustness of the link between regimes, stability and conflict can be tested by gradually increasing the intermittency threshold, and thereby more and more excluding those onsets that are likely to be part of the conflict trap.

Data Structure and Econometric Model

The data structure used in this analysis closely follows that of Raknerud & Hegre (1997) and Hegre et al. (2001). The data structure has been presented

earlier in this article, but a short repetition is warranted. We start out with a list of conflict onsets, w, and sort them ascending. At each point in time an onset occurs, t_w, we take a cross-sectional snapshot of all countries independent at that time. The dependent variable is binary, where "1" denotes that the observation is the conflict that started on that specific day, and where "0" denotes that the observation belongs to the control group. The robustness checks consist of a series of filters that creates different subsets of w denoted $w^{\delta,\varepsilon}$, where δ indicates the threshold level required as one of the values {1,2,3} and ε is one of the values {2,…,9} defining the minimum requirement of peaceful period preceding an onset between two previously active parties. Each subset $w^{\delta,\varepsilon}$ is then treated as w and each element of these subsets constitutes a cross-sectional snapshot. This can be easily done in most statistical packages through if-clauses or other subset selection mechanisms.

Analyses of binary variables, such as conflict onset, often apply a logit or probit regression model. Raknerud & Hegre (1997) argue that these models assume that the probability of a civil war, when all explanatory variables are accounted for, is constant over time. This can perhaps be remedied by temporal dummy variables, but since these temporal dummy variables are of no theoretical interest, it would be better and more efficient not to estimate them. Raknerud & Hegre (1997) shows that the semi-parametric Cox model (Cox, 1972) is unaffected by temporal variations in global conflict propensity. This model, which most often is used to model the duration of spells, can also be used in the study of onsets.

The Cox model estimates a quantity referred to as the instantaneous hazard rate and denoted $\lambda_c(t)$, for all countries c in the system at time t. In most applications, the interpretation of the hazard rate is the failure rate in the short time period $(t, t+\Delta t)$, conditional upon survival up until time t. In this analysis, the time of observation is a very short period of time, in which one country experiences a civil war onset, and where all other observations are censored. Being censored in this setting means that we know that these observations did not have an onset on that particular day, but we do not know whether they will

---

[2] See Gleditsch, Hegre & Strand (2006) for a review.

have an onset in the future (see Box-Steffensmeier & Jones, 2004: 47ff for a very good introduction). Since we know that one conflict started on the day of observation, we can interpret the hazard rate as an indicator of the risk each country has of experience that particular onset. Formally the hazard rate is defined as:

$$\lambda_c(t) = \lambda_0(t)\exp\left(\sum_{j=1}^{p} \beta_j X_{jc}(t)\right)$$

This is the product between an overall, global hazard present at time t, referred to as the baseline hazard rate and denoted $\lambda_0(t)$, and the aspects of the country in question that either increase or decrease this hazard. p is the number of explanatory variables, X_{jc} is an explanatory variable j observed for each country c, and β_j is the corresponding coefficient.

Raknerud & Hegre (1997: 389) shows that the hazard rate can be seen as an approximation of the probability of an onset at time t. Given that we know there is an onset in one country at a point in time t, the probability of that war occurring in country A is given by:

$$\Pr(\text{onset in country } A \mid \text{onset happens at } t) = \frac{\exp\left(\sum_{j=1}^{p} \beta_j X_{jA}(t)\right)}{\sum_{i \in R_t} \exp\left(\sum_{j=1}^{p} \beta_j X_{ji}(t)\right)}$$

where R_t is the set of all countries at risk of experiencing an onset (i.e. independent) at time t.

In order to understand the causes of conflict, we are more interested in the way the coefficients β_j affect the hazard rates than in the variations in the baseline hazard, which we can view as a time-specific constant term. The difference between two rates is best understood as a ratio. If we think of two countries a and b, which are similar except that a is one unit higher on the last independent variable, then the ratio between them becomes:

$$\frac{\lambda_a(t)}{\lambda_b(t)} = \frac{\lambda_0(t)\exp\left(\sum_{j=1}^{p-1}\left(\beta_j X_j(t)\right) + \beta_p X_{pa}(t)\right)}{\lambda_0(t)\exp\left(\sum_{j=1}^{p-1}\left(\beta_j X_j(t)\right) + \beta_p X_{pb}(t)\right)} = \frac{\exp\left(\beta_p X_{pa}(t)\right)}{\exp\left(\beta_p X_{pb}(t)\right)} = \exp\left(\beta_p\left(X_{pa}(t) - X_{pb}(t)\right)\right)$$

This amounts to

$$\ln\left(\frac{\lambda_a(t)}{\lambda_b(t)}\right) = \beta_p\left(X_{pa}(t) - X_{pb}(t)\right)$$

Raknerud & Hegre (1997:390) interprets the coefficient β_p as the log of the relative risk between two countries identical but for the one unit difference in X_p. The interpretation of log-relative risks is identical to the interpretation of log-odds ratios when the event in question is sufficiently rare.

Variables

Following Hegre et al. (2001) I operationalize political instability as proximity to regime change through a decay function, $f(t) = 2^{-t/\alpha}$, where t represents the number of days since the previous regime change and α represents the half-life parameter. The value of α tells us how long it takes before the effect of a preceding regime change on the risk of conflict is halved. Hegre et al. (2001: 37) operationalizes their decay function with a half-life parameter of t=1 year and base the definition of a regime change on the Polity project. The operationalization used here is based on the Gates et al. (2006) definition of regime change. It is therefore not necessarily of significance that a half-life parameter of t=2.9 years provides the best fit, even though this indicates a more persistent effect than that reported by Hegre et al.

Hegre et al. (2001) used the Polity dataset as an indicator of political regimes. The Polity coding criteria classify armed conflict as a semi-democratic trait, which introduces an endogeneity bias (Strand, 2006). Therefore I use the alternative SIP measure, developed by Gates et al. (2006). The SIP measure is included both as a linear and squared term in order to capture the curve-linear relationship described by Hegre et al. (2001).

The most important control variable is proximity to previous conflict, which is operationalized with a half-life parameter of 7 years, again in contrast

to the 16 years reported by Hegre et al. (2001: 37). The results are fairly robust for various values of α, but 7 years provided the best fit and is therefore used.

Since there can be more than one conflict in a country at a given time (Gleditsch et al., 2002: 620) proximity to previous conflict can be misleading. This variable only measures proximity to the most recent conflict termination. If there are several ongoing conflicts at the time of observation, this is not reflected in the proximity variable. Another variable, indicating the presence of an active armed conflict in the country at the time of observation supplements the proximity variable, and together these two control for past and present conflict.

Following Hegre et al. (2001) controls for spatial contagion, regime characteristics, economic development, size of country and ethnic composition are included. Economic development and country size are operationalized as the natural log of GDP/capita and population, respectively. Both indicators are based on data from Gleditsch (2002). Ethnic composition is based on Roeder's (2001) ELF data. Both linear and squared terms[3] are included, to proxy both fractionalization and polarization (see Schneider & Wiesehomeier (2006) for a thorough discussion). The spatial contagion is operationalized as the proportion of neighboring countries experiencing ongoing conflict to the total number of neighbors, as defined by Gleditsch and Ward (1999).

The conflict specific variables are operationalized at each intensity level. When a given level of δ is used to restrict the sample, a corresponding level of δ is used to define proximity to previous conflict, spatial contagion and ongoing conflict. The codebook describes these operationalizations in detail.

Analysis

Model 1 is based on the most inclusive definition of onset, including all conflicts that re-emerge as active after a minimum intermittency period of ε=2 years, and Models 2 and 3 in Table 2 are specified with ε=4 and ε=8 year rules. Proximity to regime change has a strong effect on the risk of onset, and this effect is

[3] The ELF and SIP variables are centered at its mean in order to avoid high collinearity.

associated with z-scores consistently above 1.9, which is a good sign of robustness. The effect is always statistically significant at the 5% level with a one-sided test. The inverted U-curve relationship between political regimes and conflict onset does not appear equally robust. The coefficients vary across the different values of ε, and the z-scores are hovering around the 5% level of significance for a one-sided test. It is also unclear why the effect only appears strong with ε=4 years. One might speculate that this can be due to more serious commitment problems for semi-democracies in post-conflict situations, but the positive result appears contingent on the inclusion of only a small number of observations, which points in the direction of a more basic lack of robustness.

Proximity to previous conflict is also persistently strong and robust. The decline in the coefficient is not surprising considering that altering the δ parameter excludes those cases that contribute most to this effect. Whether a country has an ongoing conflict is not associated with an average change in the risk of a new onset. The other control variables perform more or less as expected. Ethnically homogeneous countries have a lower likelihood of conflict than both polarized and fractionalized countries. Poor and populous countries are also at greater risk than small and rich countries.

Table 2 about here

Altering the casualty threshold reveals similar signs of robustness. The coefficient for Proximity to regime change is quite stable and indicates that a country with a very recent experience of regime change is 80-100% more likely to experience a conflict onset than a similar country with no recent history of institutional change. The z-scores vary between 1.77 and 2.15, indicating an overall satisfactory significance in a one-sided test. Excluding the minor conflict from the analysis strengthens the case for semi-democracy as a cause of conflict, but the relationship is not robust. The coefficients in Table 3 are much stronger, but they still vary quite a lot between the different specifications of ε and δ.

Moving from Model 1 to Model 9 we have excluded more than half of the initial set of observations. The initial analysis in Model 1 began with 212 onsets,

and the final analysis in Model 9 included only 100. Yet several variables are always strong and robust. The coefficient for Proximity to previous conflict implies that even when excluding reoccurring conflicts, war-torn countries have a much higher risk of experiencing a new conflict. Economic development and population size are also always robust predictors of conflict. The coefficients change remarkably little between the various models, indicating a rock solid relationship. Ethno-linguistic fractionalization is always significant, but the various models question whether the linear or curvilinear relationship is the better operationalization. This puzzle, however, is outside the scope of this article.

Table 3 about here

Concluding Remarks

This article has introduced a new conflict onset dataset which facilitates robustness tests. The dataset draws on the Uppsala/PRIO Armed Conflict Dataset, and extends the ACD definition through opening up aspects of the definition to the user's discression. A re-analysis of Hegre et al. (2001) shows that their finding regarding the effect of proximity to regime change is robust while the effect of semi-democratic regimes is not very robust.

Future advances in Peace Science depend on rigorous methodology. We need more precise theories and better explanations for why armed conflict is still among mankind's worst problems. Our theoretical shortcomings force us to make several arbitrary coding decisions to separate war from peace and identify conflict onsets. In order to agree on new insights, we must be certain that our advances are real advances and not unfortunate combinations of arbitrary coding decisions. The framework for robustness testing presented here can help us as a profession avoid such pitfalls, and in the end advance policy advice which adheres to Mack's (2002) call for consistency.

Bibliography

Collier, Paul, and Anke Hoeffler, 2004. "Greed and grievance in civil war." *Oxford Economic Papers* 56 (4): 563–95.

Collier, Paul, Lani Elliott, Håvard Hegre, Anke Hoeffler, Marta Reynal-Querol, Nicholas Sambanis, 2003. *Breaking the Conflict Trap. Civil War and Development Policy."* Oxford: Oxford University Press.

Cunningham, David E., Kristian Gleditsch, and Idean Salehyan, 2005. "Dyadic Interactions and Civil War Duration." Presented at the 46th Annual International Studies Association Convention, Honolulu, Hawaii. March 1-5.

Curtis, Richard, Ben Elton, Rowan Atkinson, and John Lloyd, 1998. *Blackadder. The Whole Damn Dynasty*. London: Penguin.

Fearon, James D., and David D. Laitin, 2003. "Ethnicity, insurgency, and civil war.", *American Political Science Review* 97 (1): 75-90.

Gates, Scott G., and Håvard Strand, 2006. "Modelling the Duration of Civil Wars: Measurement and Estimation Issues.", unpublished ms.

Gates, Scott, Håvard Hegre, Mark P. Jones, and Håvard Strand. 2005. "Institutional Inconsistency and Political Instability: Persistence and Change in Political Systems Revisited, 1800–1998." *American Journal of Political Science* 50(4): 893–908.

Gleditsch, Kristian S., 2002. 'Expanded Trade and GDP Data', *Journal of Conflict Resolution* 46(5): 712–724.

Gleditsch, Kristian S., and Michael D. Ward, 1999. "Interstate System Membership: A Revised List of the Independent States since 1816." *International Interactions* 25:393–413.

Gleditsch, Kristian S., and Michael D. Ward, 2001. "Measuring Space: A Minimum Distance Database." *Journal of Peace Research* 38:749-68.

Gleditsch, Kristian Skrede, 2004. "A Revised List of Wars Between and Within Independent States, 1816-2002" *International Interactions* 30:231-262.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Håvard Strand, 2002. "Armed conflict 1946-2001: A new dataset.", *Journal of Peace Research* 39 (5): 615–37.

Harbom, Lotta, 2005. *States in Armed Conflict 2004*, Research Report No 71. Uppsala: Universitetstryckeriet

Harbom, Lotta, and Peter Wallensteen, 2005. "Armed Conflict and Its International Dimensions, 1946–2004", *Journal of Peace Research* 42 (5): 623–35.

Harbom, Lotta, Stina Högbladh, and Peter Wallensteen, 2006. "Armed Conflict and Peace Agreements", *Journal of Peace Research,* 43(5): 617–631.

Hegre, Håvard, and Nicholas Sambanis, 2006. "Sensitivity Analysis of the Empirical Literature on Civil War Onset", *Journal of Conflict Resolution* 50(4): 508–535.

Hegre, Håvard, Scott Gates, Nils Petter Gleditsch and Tanja Ellingsen, 2001. "Toward a Democratic Civil Peace? Democracy, Political Change and Civil War 1816–1992", *American Political Science Review* 95(1): 33–48.

Mack, Andrew, 2002. "Civil War: Academic Research and the Policy Community", *Journal of Peace Research* 39(5): 515–525.

Marshall, Monty G., and Ted Robert Gurr, 2005. "Peace and Conflict 2005: A Global Survey of Armed Conflicts, Self-Determination Movements, and Democracy." Center for International Development & Conflict Management, Department of Government and Politics, University of Maryland

McLaughlin, Sara, Gates, Scott, Hegre, Håvard, Gissinger, Ranveig, & Gleditsch, Nils Petter. (1998). "Timing the changes in political structures: A New Polity Database." *Journal of Conflict Resolution*, 42 (2), 231–42.

Pfetsch, Frank R., and Christoph Rohloff, 2000. "KOSIMO: A Databank on Political Conflict." *Journal of Peace Research* 37 (3): 379–89.

Raknerud, Arvid and Håvard Hegre, 1997. "The Hazard of War. Reassessing the Evidence of the Democratic Peace", Journal of Peace Research 34(4): 385–404.

Roeder, Philip G., 2001. "Ethnolinguistic Fractionalization (ELF) Indices, 1961 and 1985." February 16. <http//:weber.ucsd.edu\~proeder\elf.htm>. (downloaded 15 January 2006).

Sambanis, Nicholas, 2004. "What Is Civil War? Conceptual and Empirical Complexities of an Operational Definition." *Journal of Conflict Resolution* 48 (6):814–858.

Sarkees, Meredith Reid, 2000. "The Correlates of War data on war: An update to 1997". *Conflict Management and Peace Science* 18 (1): 123–44.

Small, Melvin, and J. David Singer, 1982. *Resort to arms: International and civil war, 1816-1980*. Beverly Hills, CA: Sage.

Strand, Håvard, 2006. "Retreating From a Civil Democratic Peace? Revisiting the Relationship Between Political Institutions and Civil War." unpublished ms.

Appendix A: A List of Conflict Onsets

All conflicts listed in this Appendix are also listed in Harbom & Wallensteen (2005). The Start Date variable was collected by Gates & Strand (2006). 'Country Number' and 'Country Name' refer to the International System Membership list (Gleditsch & Ward, 1999). Only conflicts in independents states are included. 'Start Date' and 'Conflict ID' are both adopted from Gates & Strand. The Conflict ID is not directly compatible with the Conflict ID variable defined by Gleditsch et al. (2002). In order to handle various problems, a 'Sub-ID' was defined in order to separate between different stages of the conflict. Some of these, such as the presence of international involvement in a civil war, are irrelevant to this project and to Gates & Strand (2006). I have therefore recoded the ID variable so that the three first digits remain compatible with the

ACD variable, whereas the last digit is not. The original ID can be obtained by the following formula: Original ID = INT('Conflict ID'/10)*10.

Intermittent Period is only defined for the onsets that restart previously active conflicts, and is therefore listed as missing for all other conflicts. In the dataset, these observations are replaced with the value 36525, which is the average number of days in a 10-year period. This is done in order to ease the implementation of the robustness test.

Incompatibility is defined as either 'Government' or 'Territory', and gathered directly from the 2005 version of the ACD dataset (Harbom & Wallensteen, 2005). Since the intermittent period concept is not implemented in the ACD, the consequential reinterpretation of the intermediate category had to take place within this project. The effect of this reinterpretation was that some conflict periods where downgraded from intermediate (intensity level 2) to minor. Examples include the 1966–68 conflict between KDPI and the government of Iran, and some of the Baltic conflicts which were direct continuations of World War II.

The intermittent period listed in Table A–1 is based on all onsets. When running regressions on threshold-based subsets, the intermittent variables must be adjusted to disregard preceding periods of low intensity. Cuba serves as an example. For researchers exclusively interested in civil wars, the precursor to the Cuban war which took place in Santiago on July 26, 1953 should probably be disregarded. Intermittent period variables corresponding to threshold restrictions $C_{w^{\{2,3\},\varepsilon}}$ are included and labeled in the dataset.

Country Number	Country Name	Start Date	Conflict ID (ACD)	Intermittent Period	Incompatibility	Intensity
710	China	01-jan-46	1030		Government	3
630	Iran	01-jan-46	1060		Territory	1
630	Iran	01-jan-46	1070		Territory	1
365	Russia (Soviet Union)	01-jan-46	1110		Territory	1
365	Russia (Soviet Union)	01-jan-46	1120		Territory	1
365	Russia (Soviet Union)	01-jan-46	1130		Territory	3
365	Russia (Soviet Union)	01-jan-46	1140		Territory	3
350	Greece	01-mar-46	1040		Government	3
840	Philippines	04-jul-46	1100		Government	3
145	Bolivia	18-jul-46	1010		Government	3
710	China	28-feb-47	1180		Territory	3
150	Paraguay	07-mar-47	1220		Government	3
775	Myanmar	04-jan-48	1250		Territory	2

775	Myanmar	04-jan-48	1260		Territory	1
678	Yemen (Arab Republic of Yemen)	01-feb-48	1330		Government	3
94	Costa Rica	13-mar-48	1270		Government	3
775	Myanmar	28-mar-48	1240		Government	3
732	Korea, Republic of	15-aug-48	1320		Government	3
750	India	18-sep-48	1290		Government	3
666	Israel	01-jan-49	1370		Territory	2
775	Myanmar	31-jan-49	1230		Territory	3
90	Guatemala	16-jul-49	1360		Government	1
775	Myanmar	01-nov-49	1340		Territory	1
850	Indonesia	13-jul-50	1400		Territory	3
710	China	07-okt-50	1390		Territory	1
800	Thailand	30-jun-51	1430		Government	1
145	Bolivia	09-apr-52	1011		Government	1
40	Cuba	26-jul-53	1450		Government	1
850	Indonesia	20-sep-53	1460		Government	3
150	Paraguay	05-mai-54	1221		Government	1
90	Guatemala	27-jun-54	1361		Government	1
160	Argentina	16-jan-55	1500		Government	1
817	Vietnam, Republic of	01-apr-55	1520		Territory	3
750	India	01-jan-56	1540		Territory	1
710	China	01-mai-56	1390	2031	Territory	3
40	Cuba	02-des-56	1450	1225	Government	3
775	Myanmar	01-jan-57	1560		Territory	1
698	Muscat and Oman, United Kingdom	01-jul-57	1610		Territory	1
820	Malaysia	01-jan-58	1640		Government	1
850	Indonesia	15-feb-58	1460	1507	Government	2
660	Lebanon	15-mai-58	1630		Government	3
645	Iraq	14-jul-58	1620		Government	1
645	Iraq	08-mar-59	1621		Government	1
710	China	10-mar-59	1390	799	Territory	3
812	Laos	12-nov-59	1650		Government	3
775	Myanmar	30-nov-59	1670		Territory	3
490	Congo, Democratic Republic of (Zaire)	01-aug-60	1680		Territory	1
490	Congo, Democratic Republic of (Zaire)	01-sep-60	1690		Territory	1
530	Ethiopia	14-des-60	1700		Government	1
790	Nepal	15-des-60	1720		Government	1
40	Cuba	17-apr-61	1451		Government	1
220	France	22-apr-61	1730		Government	3
645	Iraq	01-des-61	1740		Territory	3
775	Myanmar	31-des-61	1341		Territory	3
530	Ethiopia	01-jan-62	1780		Territory	3
101	Venezuela	02-jun-62	1800		Government	1
678	Yemen (Arab Republic of Yemen)	26-sep-62	1331		Government	3
820	Malaysia	01-jan-63	1830		Territory	1
625	Sudan	01-jan-63	1850		Territory	3
160	Argentina	02-apr-63	1501		Government	1
645	Iraq	18-nov-63	1622		Government	1
481	Gabon	18-feb-64	1870		Government	1
490	Congo, Democratic Republic of (Zaire)	01-mai-64	1860		Government	3
42	Dominican Republic	24-apr-65	1930		Government	1
850	Indonesia	28-jul-65	1940		Territory	1

135	Peru	01-okt-65	1950		Government	1
516	Burundi	18-okt-65	1900		Government	1
483	Chad	01-nov-65	1910		Government	3
475	Nigeria	15-jan-66	2000		Government	1
652	Syria	23-feb-66	2020		Government	1
452	Ghana	24-feb-66	1980		Government	1
100	Colombia	16-aug-66	1920		Government	3
560	South Africa	26-aug-66	2010		Territory	3
750	India	01-sep-66	1990		Territory	1
90	Guatemala	01-okt-66	1362		Government	3
630	Iran	01-mar-67	1061		Territory	1
145	Bolivia	23-mar-67	1012		Government	1
811	Cambodia	02-apr-67	2030		Government	3
750	India	25-mai-67	1291		Government	1
475	Nigeria	06-jul-67	2070		Territory	3
490	Congo, Democratic Republic of (Zaire)	11-jul-67	1861		Government	1
840	Philippines	20-aug-70	2120		Territory	3
438	Guinea	22-nov-70	2110		Government	1
200	United Kingdom	01-jan-71	2190		Territory	1
580	Madagascar	01-jan-71	2140		Government	1
500	Uganda	25-jan-71	2180		Government	1
770	Pakistan	26-mar-71	2160		Territory	3
780	Sri Lanka	30-apr-71	2170		Government	3
600	Morocco	10-jul-71	2150		Government	1
625	Sudan	19-jul-71	2130		Government	1
698	Oman	01-jan-72	2210		Government	1
552	Zimbabwe (Rhodesia)	01-jan-72	2220		Government	3
165	Uruguay	28-jan-72	2230		Government	1
92	El Salvador	25-mar-72	2200		Government	1
500	Uganda	17-sep-72	2181		Government	1
840	Philippines	21-sep-72	1101		Government	3
645	Iraq	01-jan-73	1740	1027	Territory	3
160	Argentina	01-mar-73	1502		Government	3
155	Chile	11-sep-73	2250		Government	1
770	Pakistan	01-jan-74	2290		Territory	3
820	Malaysia	01-jan-74	1641		Government	1
500	Uganda	23-mar-74	2182		Government	3
800	Thailand	01-okt-74	1431		Government	2
600	Morocco	20-jan-75	2350		Territory	3
660	Lebanon	13-apr-75	1631		Government	3
771	Bangladesh	15-aug-75	2260		Territory	2
540	Angola	11-nov-75	2310		Government	3
850	Indonesia	31-des-75	2340		Territory	3
530	Ethiopia	01-jan-76	1701		Government	3
775	Myanmar	01-jan-76	1670	1827	Territory	2
850	Indonesia	01-feb-76	1940	2265	Territory	3
625	Sudan	02-jul-76	2131		Government	1
530	Ethiopia	01-sep-76	2330		Territory	3
541	Mozambique	01-jan-77	2360		Government	3
490	Congo, Democratic Republic of (Zaire)	17-aug-77	1862		Government	2
750	India	01-jan-78	2390		Territory	1
93	Nicaragua	02-feb-78	2400		Government	3

520	Somalia	09-apr-78	2410		Government	1
700	Afghanistan	27-apr-78	2370		Government	3
811	Cambodia	03-des-78	2030	1326	Government	3
630	Iran	01-jan-79	2440		Government	3
630	Iran	15-mar-79	1062		Territory	3
652	Syria	16-jun-79	2021		Government	3
411	Equatorial Guinea	03-aug-79	2430		Government	1
92	El Salvador	01-okt-79	2201		Government	3
630	Iran	03-okt-79	2450		Territory	1
670	Saudi Arabia	05-des-79	2460		Government	1
230	Spain	01-jan-80	2480		Territory	1
616	Tunisia	27-jan-80	2490		Government	1
450	Liberia	12-apr-80	2470		Government	1
678	Yemen (Arab Republic of Yemen)	01-mai-80	1332		Government	1
820	Malaysia	01-jan-81	1641	1828	Government	1
560	South Africa	01-jan-81	2510		Government	1
420	Gambia	30-jul-81	2500		Government	1
520	Somalia	01-okt-81	2411		Government	3
93	Nicaragua	01-des-81	2400	868	Government	3
452	Ghana	31-des-81	1981		Government	1
750	India	31-jul-82	2530		Territory	1
645	Iraq	01-aug-82	1623		Government	1
501	Kenya	01-aug-82	2540		Government	1
135	Peru	22-aug-82	1951		Government	3
625	Sudan	16-mai-83	1851		Territory	3
780	Sri Lanka	01-jul-83	2580		Territory	3
750	India	20-aug-83	2570		Territory	3
471	Cameroon	06-apr-84	2590		Government	1
640	Turkey	15-aug-84	2600		Territory	3
630	Iran	01-jan-86	2440	1097	Government	2
680	Yemen, Peoples Republic of	13-jan-86	2650		Government	3
115	Surinam	01-sep-86	2630		Government	1
461	Togo	23-sep-86	2640		Government	1
230	Spain	01-jan-87	2480	1827	Territory	1
645	Iraq	01-jan-87	1623	731	Government	1
439	Burkina Faso	15-okt-87	2660		Government	1
530	Ethiopia	01-jan-89	3220		Territory	1
530	Ethiopia	01-jan-89	2690		Territory	1
780	Sri Lanka	01-feb-89	2171		Government	3
150	Paraguay	03-feb-89	1222		Government	1
750	India	16-mar-89	3310		Territory	1
41	Haiti	11-apr-89	2880		Government	1
95	Panama	03-okt-89	2740		Government	1
581	Comoros	27-nov-89	2680		Government	1
910	Papua New Guinea	01-des-89	2760		Territory	1
812	Laos	01-des-89	1651		Government	1
750	India	11-des-89	2700		Territory	3
360	Romania	23-des-89	2770		Government	1
450	Liberia	30-des-89	2471		Government	3
365	Russia (Soviet Union)	19-jan-90	2840		Territory	1
365	Russia (Soviet Union)	19-jan-90	2830		Territory	1
775	Myanmar	01-mar-90	1261		Territory	1

750	India	29-mai-90	2710		Territory	1
770	Pakistan	01-jun-90	3120		Government	1
433	Senegal	01-jun-90	2820		Territory	2
432	Mali	01-jun-90	2790		Territory	1
52	Trinidad and Tobago	27-jul-90	2850		Government	1
850	Indonesia	08-sep-90	2720		Territory	1
517	Rwanda	01-okt-90	2810		Government	3
750	India	10-okt-90	1292		Government	1
645	Iraq	01-jan-91	1623	1097	Government	3
630	Iran	01-jan-91	2440	731	Government	2
230	Spain	01-jan-91	2480	1097	Territory	1
540	Angola	01-jan-91	2940		Territory	1
41	Haiti	07-jan-91	2881		Government	1
451	Sierra Leone	01-apr-91	2890		Government	3
345	Yugoslavia (Serbia)	26-jun-91	2920		Territory	3
345	Yugoslavia (Serbia)	27-jun-91	2910		Territory	1
640	Turkey	13-jul-91	2900		Government	1
775	Myanmar	15-okt-91	1240	1122	Government	1
522	Djibouti	12-nov-91	2860		Government	1
516	Burundi	26-nov-91	1901		Government	3
461	Togo	28-nov-91	2641		Government	1
615	Algeria	01-des-91	2930		Government	3
372	Georgia	22-des-91	2870		Government	1
775	Myanmar	29-des-91	1250	1197	Territory	2
750	India	01-jan-92	2530	1096	Territory	1
850	Indonesia	01-jan-92	2340	731	Territory	2
344	Croatia	01-jan-92	2970		Territory	1
372	Georgia	01-jan-92	3000		Territory	1
373	Azerbaijan	01-jan-92	2950		Territory	3
775	Myanmar	01-jan-92	1561		Territory	1
101	Venezuela	04-feb-92	1801		Government	1
359	Moldova	01-mar-92	3010		Territory	1
346	Bosnia-Herzegovina	07-apr-92	2960		Territory	3
702	Tajikistan	29-jun-92	3020		Government	3
750	India	31-jul-92	1541		Territory	1
372	Georgia	14-aug-92	2990		Territory	3
436	Niger	01-okt-92	2800		Territory	1
750	India	12-okt-92	2390	1522	Territory	1
750	India	01-jan-93	3310	732	Territory	1
775	Myanmar	01-jan-93	1670	1566	Territory	3
840	Philippines	01-jan-93	2120	732	Territory	3
651	Egypt	10-mar-93	2980		Government	1
346	Bosnia-Herzegovina	01-apr-93	3050		Territory	3
373	Azerbaijan	04-jun-93	3030		Government	1
630	Iran	01-jul-93	1062	1035	Territory	2
346	Bosnia-Herzegovina	03-okt-93	3040		Territory	1
365	Russia (Soviet Union)	03-okt-93	3070		Government	1
484	Congo	03-nov-93	3170		Government	1
70	Mexico	01-jan-94	3080		Government	1
540	Angola	01-jan-94	2940	732	Territory	1
750	India	01-jan-94	2710	746	Territory	1
432	Mali	01-jan-94	2790	1097	Territory	1

500	Uganda	01-jan-94	2183		Government	3
678	Yemen (Arab Republic of Yemen)	27-apr-94	3100		Territory	3
365	Russia (Soviet Union)	11-des-94	3090		Territory	3
775	Myanmar	01-jan-95	1230	946	Territory	2
770	Pakistan	01-jan-95	3120	1462	Government	1
373	Azerbaijan	15-mar-95	3031		Government	1
645	Iraq	01-jan-96	1740	731	Territory	2
775	Myanmar	01-jan-96	1561	1096	Territory	1
530	Ethiopia	01-jan-96	2331		Territory	1
775	Myanmar	01-jan-96	1261	1827	Territory	1
436	Niger	01-feb-96	3150		Territory	1
530	Ethiopia	26-apr-96	2690	1795	Territory	1
70	Mexico	01-jul-96	3081		Government	1
790	Nepal	13-jul-96	1721		Government	3
630	Iran	27-jul-96	1062	939	Territory	2
530	Ethiopia	10-aug-96	3140		Territory	1
490	Congo, Democratic Republic of (Zaire)	17-okt-96	1863		Government	3
850	Indonesia	01-jan-97	2340	1462	Territory	2
630	Iran	01-jan-97	2440	1097	Government	2
517	Rwanda	01-jan-97	2810	897	Government	3
483	Chad	01-jan-97	1910	732	Government	2
775	Myanmar	16-mar-97	3320		Territory	1
531	Eritrea	01-mai-97	2300		Government	1
484	Congo	05-jun-97	3170	1221	Government	3
436	Niger	01-sep-97	2800	975	Territory	1
581	Comoros	05-sep-97	3160		Territory	1
345	Yugoslavia (Serbia)	28-feb-98	3210		Territory	3
404	Guinea-Bissau	07-jun-98	3190		Government	3
200	United Kingdom	15-aug-98	2190	2419	Territory	1
540	Angola	01-sep-98	2310	975	Government	3
570	Lesotho	23-sep-98	3200		Government	1
522	Djibouti	01-jan-99	2860	1467	Government	1
850	Indonesia	01-jan-99	2720	2742	Territory	2
530	Ethiopia	01-jan-99	3220	2775	Territory	1
365	Russia (Soviet Union)	07-aug-99	3090	1087	Territory	3
365	Russia (Soviet Union)	10-aug-99	3230		Territory	1
750	India	01-jan-00	1541	893	Territory	1
450	Liberia	01-jan-00	2472		Government	3
704	Uzbekistan	17-aug-00	3240		Government	1
438	Guinea	01-sep-00	2111		Government	1
520	Somalia	01-jan-01	2411	1522	Government	1
343	Macedonia (FYROM)	22-jan-01	3260		Government	1
482	Central African Republic	28-mai-01	3250		Government	1
540	Angola	01-jan-02	2940	1097	Territory	1
484	Congo	01-apr-02	3170	867	Government	2
437	Ivory Coast	19-sep-02	3280		Government	1
531	Eritrea	01-jan-03	2300	1097	Government	1
750	India	01-jan-03	2530	731	Territory	1
625	Sudan	09-apr-03	2132		Government	3
645	Iraq	01-jan-04	1624		Government	3
41	Haiti	05-feb-04	2882		Government	1
704	Uzbekistan	01-mar-04	3240	1156	Government	1

372	Georgia	01-aug-04	3001		Territory	1
475	Nigeria	01-sep-04	3330		Territory	1

Table 2: Cox Estimates of Conflict Onset Based On Most Inclusive Casualty Threshold.

	Model 2	Model 3 δ=1	Model 4
	ε= 2 yr.	ε= 4 yr.	ε= 8 yr.
Ethno-Linguistic Frac.	2.310***	2.192**	2.096**
	(2.646)	(2.356)	(2.178)
Ethno-Linguistic Frac.^{2}	0.049***	0.052**	0.095*
	(-2.580)	(-2.381)	(-1.844)
Economic Development	0.674***	0.611***	0.611***
	(-4.096)	(-4.670)	(-4.437)
Population	1.328***	1.324***	1.281***
	(5.840)	(5.405)	(4.536)
SIP	0.942	1.193	1.244
	(-0.264)	(0.721)	(0.853)
SIP^2	0.194	0.115**	0.155*
	(-1.640)	(-2.004)	(-1.658)
Ongoing Conflict	1.000	0.831	0.789
	(-0.001)	(-0.936)	(-1.109)
Spatial Lag of Conflict	0.954	0.914	0.873
	(-0.150)	(-0.262)	(-0.375)
Proximity to Regime Change	1.566*	1.631*	2.018***
α=2.9 years	(1.904)	(1.929)	(2.701)
Proximity to Conflict	3.583***	2.212***	1.765**
α=7 years	(6.317)	(3.495)	(2.329)
Log Likelihood	-934.93	-812.36	-748.37
Log Likelihood null model	-1032.74	-879.45	-800.60
N	31276	26732	24330
Number of Countries	165	165	165
Number of Civil Wars	212	181	165

Table 3: Cox Estimates of Conflict Onset Based On Alternative Casualty Thresholds.

	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
	δ=2			δ=3		
	ε= 2 yr.	ε= 4 yr.	ε= 8 yr.	ε= 2 yr.	ε= 4 yr.	ε= 8 yr.
Ethno-Linguistic Frac.	2.147*	2.102*	2.045	2.311**	2.521**	2.308*
	(1.867)	(1.711)	(1.639)	(1.971)	(2.052)	(1.850)
Ethno-Linguistic Frac.^{2}	0.020**	0.017**	0.026**	0.017**	0.018**	0.022**
	(-2.445)	(-2.382)	(-2.144)	(-2.467)	(-2.285)	(-2.173)
Economic Development	0.706***	0.618***	0.628***	0.706***	0.660***	0.646***
	(-2.776)	(-3.490)	(-3.308)	(-2.723)	(-2.994)	(-3.080)
Population	1.300***	1.301***	1.285***	1.277***	1.280***	1.256***
	(4.155)	(3.912)	(3.672)	(3.771)	(3.594)	(3.266)
SIP	0.662	0.890	0.943	0.605	0.739	0.791
	(-1.327)	(-0.356)	(-0.179)	(-1.557)	(-0.894)	(-0.685)
SIP^2	0.087*	0.075*	0.096	0.093*	0.052**	0.079*
	(-1.864)	(-1.832)	(-1.641)	(-1.755)	(-2.054)	(-1.742)
Ongoing Conflict	0.762	0.650	0.594*	0.820	0.703	0.686
	(-1.057)	(-1.450)	(-1.656)	(-0.728)	(-1.141)	(-1.181)
Spatial Lag of Conflict	1.053	0.942	0.933	1.018	0.944	0.907
	(0.113)	(-0.118)	(-0.134)	(0.037)	(-0.109)	(-0.179)
Proximity to Regime Change	1.857**	1.809*	2.051**	1.919**	1.803*	1.967**
α=2.9 years	(2.028)	(1.792)	(2.149)	(2.091)	(1.766)	(2.005)
Proximity to Conflict	6.898***	3.962***	3.578***	8.032***	5.340***	4.937***
α=7 years	(7.270)	(4.512)	(4.044)	(7.784)	(5.635)	(5.232)
Log Likelihood	-523.99	-460.44	-447.02	-496.89	-445.16	-432.04
Log Likelihood null model	-599.94	-510.79	-491.37	-575.86	-501.69	-482.21
N	17631	15064	14382	16867	14743	14053
Number of Countries	165	165	165	165	165	165
Number of Civil Wars	124	106	102	119	104	100