The Severity of the Colombian Conflict: Cross-Country Datasets versus New Micro Data*

Jorge A. Restrepo
Department of Economics
Universidad Javeriana
Royal Holloway College
University of London and CERAC
Jarestrepo@javeriana.edu.co

Michael Spagat (Corresponding Author)
Department of Economics
Royal Holloway College
University of London
Egham, Surrey, TW20 0EX, UK
CEPR, CERAC and WDI
M.Spagat@rhul.ac.uk

and

Juan F. Vargas
Royal Holloway College
University of London and CERAC
J.F.Vargas@rhul.ac.uk

This version: June 14, 2005

Abstract

We compare the treatment of Colombia in large cross-country conflict datasets with the information of the unique CERAC dataset on the Colombian conflict. The big datasets display a strong tendency to record fewer killings than does CERAC. Moreover, when the big datasets provide annual time series on the conflict these figures look either erratic or flat compared to CERAC's and often move in different directions. We examine in detail the criteria of the Uppsala Conflict Data Project (UCDP) for dataset inclusion and find them considerably more restrictive than CERAC's. The primary differences are that UCDP generally excludes attacks purely on

* We thank Kristine Eck, Nils Petter Gleditsch, Helena Grusell, Bethany Lacina, an anonymous referee and numerous researchers from the Uppsala Conflict Data Project for many helpful comments. We are also thankful to seminar and conference participants at Banco de la República, Royal Holloway, 8th ECAAR conference on Economics and Security, Bristol, UK, the University of Manchester and the International Studies Association Conference in Hawaii. The RSF fund of Royal Holloway and the Department of Economics contributed funding to our research project. Both Restrepo and Vargas acknowledge financial support from Banco de la República in Colombia and Vargas also acknowledges financial support from the ORS Award Scheme. The data used in the paper is downloadable from: http://www.cerac.org.co.

civilians and any activity of illegal right-wing paramilitary groups. We argue that these omissions impoverish our perception of many civil wars. We calculate a modified series based on UCDP methodology and CERAC raw information. This exercise closes on average 56% of the gap between the two approaches. The remainder appears to derive mainly from a number of small events in CERAC but not UCDP, reflecting the limits of English-language press coverage of Colombia, upon which UCDP data is based. The gap with other big datasets is also closed. The dynamics of the lower-bound UCDP curve clearly resemble the modified CERAC curve so UCDP does reasonably well on its own terms. A brief Northern Ireland case study is consistent with our Colombia conclusions. We recommend that conflict researchers should prioritize the construction of more micro datasets that will facilitate detailed studies of conflict intensity and its dynamics.

Introduction

The development of big cross-country conflict datasets has been vital for the advance of conflict studies, particularly for the influential literature based on empirical analysis of civil conflicts. Pioneering works in this field include the Correlates of War project (hereafter COW; Small and Singer, 1982), the Civil War Termination project (CWT; Licklider, 1995) and the Uppsala/PRIO dataset (Gleditsch et al., 2002).[1]

Restrepo, Spagat and Vargas (2004) introduced an entirely different approach to the study of civil conflict, developing a general methodology for the in-depth measurement of conflict activity in a single conflict. Restrepo et al. (2004) applied this approach to construct a detailed, micro-level dataset for Colombia which we will refer to as CERAC in reference to the Bogotá-based think tank that maintains the data (http://www.cerac.org.co). Restrepo et al. (2004) and Restrepo and Spagat (2004a) analyze the dataset so that its general contours are now clear. Therefore, the time is ripe to compare CERAC with the Colombia components of the large cross-country datasets.

This paper will serve two main purposes. First, we will evaluate the killing figures for Colombia used in the cross-country datasets by comparing them with those of CERAC. We show that the cross-country datasets generally produce lower figures than does CERAC. We, therefore, believe that some of the other ongoing data-collection efforts should consider adjusting their numbers upward, as long as these changes are consistent with the methodology these projects are applying to other conflicts.[2] Some of

[1] Two comprehensive references on the issues and characteristics of these datasets are Eck (2003) and the webpage for the 2001 Uppsala Conflict Data conference (www.pcr.uu.se/conferenses/euroconference/workpapers.html).

[2] Specialists on the Colombian conflict may also want to revise some of their work. Most empirical work in the field uses the homicide rate per 100,000 people as the best proxy of violence (see the survey in Riascos

the cross-country datasets give time series so we compare the dynamics of these series with CERAC dynamics. We find that some of these datasets have erratically fluctuating Colombia figures while another has rather flat dynamics compared to CERAC. Moreover, these curves often move in different directions than does the CERAC curve.

We then examine the methodology of the Uppsala Conflict Data Project (UCDP) and find that it specifically screens out many events that CERAC includes. In an attempt to compare like with like we calculate a hybrid curve that combines UCDP methodology with CERAC raw information. This modification eliminates 56% of the difference in total deaths between CERAC and UCDP's preferred lower-bound estimate. Thus, the methodological differences between the two approaches are substantial but do not explain the entire divergence. Inspection of the datasets suggests that the remainder derives from a number of small events in CERAC but not UCDP. Finally, we compare the dynamics of the modified CERAC curve and the UCDP lower bound curve and find a clear resemblance, although the upper-bound UCDP curve looks quite different. So the lower bound does a reasonable job of capturing dynamics according to Uppsala's own criteria.

We perform a similar exercise for Northern Ireland, comparing UCDP data with that of Malcolm Sutton (Sutton, 1994 and http://cain.ulst.ac.uk/sutton/). Sutton, like CERAC, applies less restrictive criteria for dataset inclusion than does UCDP. On the other hand, screening Sutton's data using UCDP criteria produces a much closer match than obtains in the CERAC-UCDP comparison. We believe this is because UCDP's computer searching of English-language sources performs much better in the English-

---

and Vargas, 2004) but Restrepo et al. (2004) show that the dynamics of the homicide rate are different from CERAC intensity dynamics.

speaking environment of Northern Ireland than in Spanish-speaking Colombia so the "small events exclusion effect" is not significant.

Our second main purpose is to provide a general quality check on the cross-country datasets. It would be impossibly difficult and expensive for a cross-country dataset to treat every single country at the level of detail and with the degree of care that CERAC applies to Colombia. Still, by comparing the big datasets with each other and with CERAC at their main point of intersection we are evaluating quality. Admittedly, Colombia is just one case. But our Colombia results are generally consistent with our findings for Northern Ireland. Moreover, the results make sense. For example, measures of battle deaths that omit attacks on civilians and illegal paramilitary activity should come out lower than measures that do and local sources should pick up more events than do international English ones. We hope that more micro-level datasets on conflict will become available in the future to enable further investigations into the quality of cross-country conflict data.

Types of conflict data

The early development of conflict databases such as COW (http://www.correlatesofwar.org/cow2%20data/WarData/IntraState/Intra-State%20War%20Format%20(V%203-0).htm) focused sensibly on compiling lists of wars, sometimes complemented by broad estimates of the total number of victims. Much empirical conflict work has worked off of this basis (e.g., Doyle and Sambanis, 2000: D&S; Fearon and Laitin, 2003: F&L, Collier & Hoeffler, 2004). More recently, much serious effort has gone into the collection of conflict intensity information, most notably in

the work of Lacina and Gleditsch (2005) and the UCDP. Lacina and Gleditsch (2005) provide total battle-death figures on all conflicts, 1946-2002, based on the best available secondary sources for each conflict. UCDP provides battle-death time series beginning in 1989 for all conflicts through the use of the newswire-scanning IDEA software (Bond et al. 2003). Some researchers, such as Mesquida and Weiner (1999) and Hartzell and Hoddie (2003) have made use of intensity information. Other intensity measures have also been used in the literature including duration (Fearon, 2004 and Collier, Hoeffler and Söderbom, 2004) and the size of the conflict area (Buhaug and Gates, 2002). In this paper we focus primarily on conflict death information, with a special interest in the time dimension to assess the potential that cross-country datasets have for monitoring conflict dynamics.

A variety of methods have been used to measure conflict deaths. Roberts et al. (2003) and Roberts et al. (2004) both use survey methods to study war-related excess deaths in the Democratic Republic of Congo and Iraq respectively. This technique involves sending interview teams to a random sample of locations to gather information on mortalities before and during a conflict. Another approach with a long history in the international relations literature is to systematically mine press records.[3] Computer technology has provided new impetus to this approach (Schrodt and Gerner, 1994). King and Lowe (2003) provide evidence that machine coding performs at least as well as human coding for the collection of typical conflict data. A third general approach is to work off of information provided by sources compiled for other purposes such as human rights reports and truth and reconciliation commissions (e.g. Brunborg, Lyngstad and Udral, 2003). Davenport and Ball (2002) compare Guatemalan conflict information gleaned from newspapers, human rights reports and interviews and find considerably

[3] See sources in King and Lowe (2003).

differing accounts by source, suggesting the need for conflict data-builders to utilize multiple sources of information. Similarly, Mueller (1995) shows that the initial press accounts of the 1991 Gulf War significantly overestimated Iraqi casualties.

We emphasize two general points about CERAC before proceeding to a general description. First, it is hand-coded, which is a virtual necessity given that it is based primarily on Spanish-language sources. The proprietary IDEA software (Bond et al., 2003) for machine coding at this stage operates exclusively in English. Such software utilizes the structure of the English language and, therefore, it would be a major project to develop such software for Spanish or for other languages. Moreover, CERAC incorporates detailed geographical information which is not possible to collect presently through the use of machine coding. Second, CERAC is primarily based on the raw information of an NGO but is not equivalent to this information, as the information is verified against and supplemented with information from a variety of other sources.

The data on Colombia

CERAC is the first time-series dataset for the Colombian civil war that is detailed (more than 21,000 events), high-frequency and long. It allows analysis of the actions of all participants in the Colombian conflict over a 17-year period. Its conflict-measurement methodology is based on events as the unit of data inclusion and analysis. For each event the database records a set of characteristics: date; geographical location; whether or not there was a clash between two or more forces or a (one-sided) attack of which we distinguish between many types; and the group(s) involved. It also includes the number of killings and injuries resulting from every event. In this way, researchers can gauge not

only the dynamics of the conflict across space and time, but also the intensity of various conflict activities.

We summarize here the main characteristics of the dataset and refer the reader to Restrepo et al. (2004) for details. The dataset is built using events listed in the annexes to the periodicals *Justicia y Paz* and *Noche y Niebla* published quarterly by the Colombian NGO'S CINEP and the Comisión Intercongregacional de Justicia y Paz (hereafter, CINEP). Most of the event information comes from primary sources. CINEP uses this information in its reports, focusing on the measurement of human rights violations, violations to international humanitarian law and political violence, connected or not with the conflict. CERAC, on the other hand, focuses on civil war dynamics. Therefore, CINEP's database organization and statistical analysis are entirely inappropriate for CERAC's purposes. Fortunately, the raw information of CINEP is so extensive that CERAC researchers are able to distil from it just its war-relevant components. Working from the detailed list of events published in the annexes to the reports, CERAC researchers identify and code events following their own criteria designed to include all conflict events and only those events. The specific inclusion criterion is that there must be clear evidence that an event was carried out by an organized, politically motivated group. A small number of events with unknown perpetrators do meet this criterion, e.g., some bombings of economic infrastructure targets which are surely perpetrated by guerrillas although the specific group may be unknown.

In the original dataset and in quarterly updates the CERAC team follows a stringent quality control regime in cleaning the data that proceeds in four stages, covering both event inclusion and the coding of events. First, they randomly sample a large

number of events and check against the CINEP source that they are properly included and coded. Second, they randomly sample events, look up these events in press archives and again verify their inclusion and coding. This is a test both of the transfer of information from the CINEP source to CERAC and of the quality of the CINEP raw information itself, which turns out to be high. Third, they find all the major events in the dataset and carefully investigate each one in the press record. Finally, they compare lists of significant events from other sources, such as Human Rights Watch and Colombian government reports, with CERAC, occasionally adding events after a thorough investigation.

International cross-sectional datasets

Table I lists the main relevant datasets, summarizes their nature and relates them to Colombia. This is not an exhaustive survey like the one provided in Eck (2003). Rather, we select influential datasets that are relatively accessible, focus on intra-state rather than inter-state conflict and use quantitative fatality thresholds.[4]

The object of study varies across the datasets both in terminology and content. COW, CWT, D&S and F&L all focus on the category "civil war". There is some variation in the definition of civil war across these datasets but at the intersection of the qualitative components of these definitions there are the following requirements: civil wars occur within the recognized boundaries of a state; the state fights against organized

[4] The dataset of the International Institute for Strategic Studies (IISS, see http://www.iiss.org/showpage.php?pageID=25) is an exceptional case that does not employ quantitative violence thresholds but we list it anyway since it provides a time series for its intensity measure. In addition, the dataset of the Stockholm International Peace Research Institute (SIPRI, see http://www.sipri.org/contents/conflict/MAC_definition.html) is widely consulted. We omit this dataset from the present paper because SIPRI works closely with the research team at Uppsala University that already contributes to two of the datasets we analyze below.

groups striving for political power; the rebels effectively challenge the sovereignty of the state in some regions; animosity between parties of the conflict together with the fact that peace would require living together affects the type of peace settlement that can be reached. It is quite clear that the Colombian conflict satisfies these criteria.[5] Other datasets categorize their object of study differently. IISS, Project Ploughshares, and the Uppsala/PRIO and UCDP datasets list "armed conflicts", while for the State Failure Task Force (SFTF) and the Third Party Intervention (TPI) Colombia is respectively a "revolutionary war" and an "ideological civil conflict".[6] Finally, the World Military and Social Expenditures yearbooks (WMSE) studies simply "wars". Nevertheless, the definitions of the various datasets share much common ground with the civil war definition given above.

Except for IISS (see footnote 4), all of the databases define violence thresholds that a conflict must cross for inclusion (Table I). CERAC figures indicate that all these thresholds are, indeed, comfortably satisfied for Colombia from 1988 (the first year in the CERAC dataset) to the present.

Beyond the range of the object of study, there remains considerable variety among datasets. Most datasets are academic projects seeking to underpin cross-country studies while others are conflict monitoring projects primarily for educational or advocacy proposes (IISS, Project Ploughshares, WMSE). Some databases are not

[5] Rabasa and Chalk (2001) gives a good overview of the Colombian conflict. Nevertheless, many analysts of the Colombian conflict insist that, although it is a very serious affair, the conflict should not be described as a civil war. Posada (2001), for example, stresses that the illegal armed groups in Colombia enjoy very little popular support. In his view, civil war terminology endows these violent actors with undeserved legitimacy, constantly encouraging the notion that the State should negotiate with them and address their concerns. In this view, conflicts should be classified as civil wars only when insurgents enjoy substantial civilian support. Such arguments are alien to Table I so we do not pursue them further here, although perhaps the quantitative civil war literature should take account of them.

[6] Lacina and Gleditsch (2005) also study "armed conflicts". We omit this dataset from our comparison because they have adopted CERAC figures.

regularly updated since they were created for specific projects that have already been completed. Other databases are updated regularly. Some datasets include time series on killing rates while others just give aggregate numbers or even omit conflict intensity numbers entirely. F&L use intensity as a screen for dataset inclusion but do not include intensity information in their dataset.

Table I. Treatment of the Colombian Conflict in Cross-country Datasets

Dataset	Source	Type of Event (a)	Violence threshold	Time Series	Start date (b)
Civil War Termination (CWT)	Licklider (1995)	Civil War	1,000 battle-related fatalities per year. The weaker side must impose casualties on its oponent of at least 5% of its own	No	1978
Correlates of War (COW) - Intrastate War	Sarkees (2000)	Civil War		No	1984
Doyle and Sambanis (D&S)	Doyle and Sambanis (2000)	Civil war	1000 battle-related fatalities per year	No	1978
Fearon and Laitin (F&L)	Fearon and Laitin (2003)	Civil War	The conflict killed or has killed as least 1000 over its course, with a yearly average of at least 100. At least 100 of the dead are on the side of the government (including civilians).	From 1945	1963
International Institute for Strategic Studies (IISS)	IISS (2004)	Internal Armed Conflict	0	From 2000	1963
Project Ploughshares	Project Ploughshares (2004)	Armed Conflict	1000 battle related fatalities during the course of the conflict	From 1998	1964
State Failure Task Force (SFTF)	Marshall, Gurr and Harff (2002)	Revolutionary War	An average of 100 fatalities per year	From 1948	1984
Third-Party Intervention (TPI)	Regan (2002)	Ideological Civil Conflict	200 in total	No	1984
Uppsala Conflict Data Project (UCDP), Uppsala University and International Peace Research Institute, Oslo (PRIO); Uppsala/PRIO	Gleditsch et al. (2002)	Armed Conflict	1) Minor Armed Conflict: at least 25 battle-related deaths per year and fewer than 1,000 battle-related deaths during the course of the conflict 2) Intermediate Armed Conflict: At least 25 battle-related deaths per year and an accumulated total of at least 1,000 deaths, but fewer than 1,000 in any given year 3) War: At least 1,000 battle-related deaths per year	From 1946	1965
Uppsala Conflict Data Project (UCDP)	UCDP (2004)	Armed Conflict		From 1989	1965
World Military and social Expenditures (WMSE)	Sivard (1991)	War	1000 battle related fatalities per year	No	1986

Notes:
(a) Different datasets claim to be dealing with these different types of conflict
(b) Starting year of the current conflict according to each dataset. They all consider the conflict as ongoing at the moment of their last update.

Generally the datasets do not attempt to build time series on killings and the few exceptions give quite wide ranges (Table II). When there are no underlying time series

we find it difficult to place great confidence in aggregate numbers. Before the 1990s only the SFTF, Uppsala/PRIO and UCDP datasets provide annual conflict intensity time series.[7] The former gives a discrete intensity index that varies from 0 (less than 100 fatalities per year) to 4 (more than 10,000 per year) with very wide ranges in between. Uppsala/PRIO, similarly, provides an index that goes from 1 (at least 25 battle-related deaths per year and fewer than 1,000 over the course of the conflict) to 3 (at least 1,000 per year), and UCDP offers a narrower range of actual figures for battle-related fatalities beginning in 1989.[8]

Comparing numbers

Table II compares CERAC with the data from the cross-country datasets that provide time series on conflict-related killing. Some of the datasets give ranges of figures so that these comparisons are sometimes ambiguous. Nevertheless, half the estimates in the table are clear underestimates compared to CERAC.[9] The degree of underestimation varies widely and is often very high. For example, the mean of the range for UCDP never exceeds 2/3 of the CERAC figures, is less than half of CERAC's figure in most years and is less than 1/3 of CERAC's figure in 2002. The SFTF dataset in 1999 accounts for the only clear overestimate in Table II. The numbers in this dataset jump

[7] The UCDP of Uppsala University has been collecting data on armed conflicts since the late 1980s. It has recently begun collaboration with the International Peace Research Institute, Oslo (PRIO) and it has expanded its database to cover the post World War II period (Gleditsch et al., 2002) producing what we call the Uppsala/PRIO dataset, the most recent update of which being Eriksson and Wallensteen, 2004. The UCDP further expanded their coverage by adding new variables and released a searchable web-based dataset on armed conflicts (http://www.pcr.uu.se/database/index.php) that we call in our tables UCDP.

[8] Despite their similar approaches SFTF and Uppsala/PRIO-UCDP treat Colombia very differently. SFTF's dataset inclusion criterion is less strict but, nevertheless, codes no conflict between 1960 and 1984 while Uppsala/PRIO and UCDP date the start of the current conflict as 1965.

[9] To save space we will consistently write that the cross-country datasets underestimate or overestimate without always including the phrase "relative to CERAC."

around erratically between 1996 and 2001 while the conflict was intensifying continuously, raising the possibility that the number for 1999 is simply an error. Even when the figures in Table II are compatible with CERAC's the ranges are very wide. Table II suggests that the big datasets generally underestimate the magnitude of the Colombian conflict, especially since the CERAC team follows a rather conservative approach, excluding events when it does not find clear evidence of conflict relatedness. For example, CERAC will generally exclude a political assassination unless there is a strong reason to believe that it was not perpetrated by common criminals and not associated with non-conflict motives such as corruption and personal vendettas.

Table II. Total Annual Deaths When Time Series are Available

year	CERAC	IISS	Ploughshares	State Failure	Uppsala/PRIO	UCDP
1988	1,859			1,000 - 5,000	25 - 1,000	
1989	1,236			1,000 - 5,000	25 - 1,000	152 - 732
1990	1,820			1,000 - 5,000	> 1,000	395 - 1,229
1991	1,860			1,000 - 5,000	25 - 1,000	578 - 1,364
1992	2,036			1,000 - 5,000	> 1,000	541 - 1,478
1993	1,560			1,000 - 5,000	> 1,000	187 - 1,331
1994	1,375			1,000 - 5,000	25 - 1,000	333 - 1,243
1995	1,330			1,000 - 5,000	25 - 1,000	324 - 1,105
1996	1,582			100 - 1,000	25 - 1,000	817 - 1,300
1997	1,741			1,000 - 5,000	25 - 1,000	467 - 1,703
1998	2,417		2,000 - 4,000	1,000 - 5,000	> 1,000	939 - 1,138
1999	2,710		2,000 - 3,000	**> 10,000**	> 1,000	827 - 1,837
2000	3,101	≤ 3,000	1,200	1,000 - 5,000	> 1,000	938 - 1,153
2001	3,245	≤ 3,000	> 2,500	100 - 1,000	> 1,000	883 - 1,362
2002	4,038	≤ 3,000	3,500		> 1,000	972 - 1,309

Underestimates are underlined
Overestimates are **bold**

SFTF, Uppsala/PRIO and UCDP classify conflicts into several intensity categories (Tables I & II). SFTF usually chooses the correct band, given CERAC's figures, but their bands are very wide. Uppsala/PRIO has very wide bands and often chooses a lower intensity category than CERAC's numbers suggest would be correct.

UCDP provides two separate annual intensity measures. The first is a classification into the same categories as in Uppsala/PRIO, so we do not include it in the table. The second are tighter estimates of annual killings which are systematically below CERAC's figures. In the next section we analyze the sources of the widespread discrepancies.

Figures 1, 2 and 3 graph CERAC's numbers against those of SFTF, Uppsala/PRIO and UCDP. The shaded part of figures 1 and 2 represent the ranges on SFTF and Uppsala/PRIO respectively and the dotted lines of figure 3 that of UCDP. First, note that the erratic fluctuation of SFTF and Uppsala/PRIO suggest that they are measuring conflict intensity poorly. Second, a key point of Restrepo et al. (2004) was that there was a major upsurge in the conflict between 1996 and 2002. It is difficult to read this characteristic into the other datasets. SFTF would suggest a huge decline in conflict intensity. The Uppsala/PRIO numbers do show an upsurge for 1994-2002 but its classifications between 1988 and 1994 fluctuate so much that one cannot, on their basis, develop much confidence that the conflict really was intensifying during 1994-2002. The upper UCDP range is essentially flat during the upsurge. The lower range does trend up during this period, although not as dramatically as the CERAC numbers. Nevertheless, UCDP places highest confidence in the lower series so there is a definite correspondence here between UCDP and CERAC. On the other hand, the distance between UCDP's upper and lower ranges narrows during the upsurge period, suggesting increasing confidence in their rather flat pattern.

Figure 1. CERAC vs. SFTF

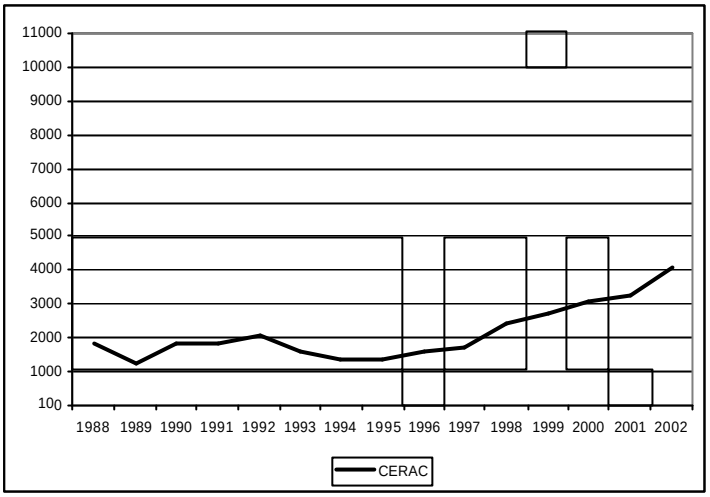


Figure 2. CERAC vs. Uppsala/PRIO

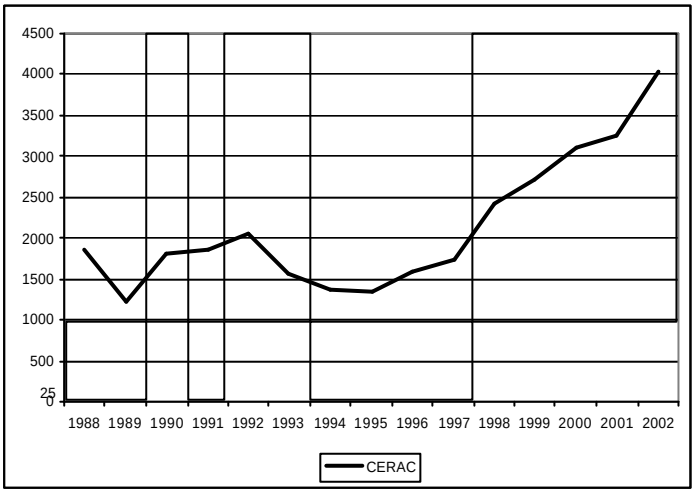
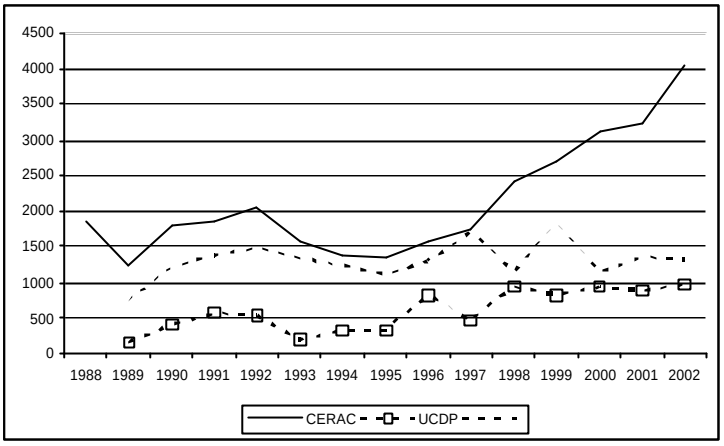


Figure 3. CERAC vs. UCDP



In Table III we compare CERAC's numbers with those in all the cross-country datasets presented in Table I. Since the different datasets cover different years in the conflict we compute annual average killing rates in each case.[10] Seven of the eleven datasets with numbers underestimate the killing rate while three datasets give overestimates and one offers ranges that include CERAC's number. Of the observable overestimations, COW and COW2 are very close to CERAC's figure while WMSE is much higher but, unlike all the other datasets including CERAC, it often includes indirect in addition to direct war-related deaths.

Table III. Annual Averages for Killing Rates

[10] When datasets give upper and lower bounds we use these to create upper and lower bounds for our annual averages in Tables III and IV. When datasets sometimes do not give lower and upper bounds but occasionally have an entry such as ">10,000" as SFTF does, we treat that number as both an upper and a lower bound.

Dataset	(1)	(2)	(3)
CERAC	1988 - 2002	31,910	2,127
CWT	1978 - 1991	15,849	1,132
COW	1984 - 1992	22,000	**2,444**
COW2*	1984 - 1997	31,000	**2,214**
D&S	1978 - 1992	22,000	1,467
IISS	1963 - 2002	56,000	1,400
Ploughshares	1964 - 2002	50,000	1,282
SFTF	1984 - 2001	22,500 - 75,000	1,250 - 4,167
TPI	1984 - 1999	11,000	688
Uppsala/PRIO	1965 - 2002	8,750 - 23,375	230 - 615
UCDP	1989 - 2002	8,353 - 18,284	597 - 1,306
WMSE	1986 - 1990	22,000	**4,400**

(1) Years of the Colombian conflict covered in each dataset.

(2) Total number of deaths during the period reported in (1)

(3) Annual average number of deaths during the period reported in (1)

* COW2 is the 1997 update of the original COW dataset (Sarkees, 2000)

Conflict intensity varies from year to year and the years of coverage of the various datasets vary as well. Therefore, the comparisons of Table III are potentially misleading. We address this issue by presenting in Table IV average death tolls for each dataset for the years of overlap with the CERAC data and compare those figures with the CERAC average death toll for these overlap years.[11] The results turn out to be identical with those of table III. We can, therefore, be confident about the relationship between the figures in each dataset and the CERAC dataset.

Table IV. Annual Averages for Killing Rates in Overlap Years

[11] Note that the figures for column 3 in Table III often coincide with those of column 3 in Table IV. This happens either when a dataset does not provide a time series or when years covered by a time series are contained in the years covered in CERAC.

Dataset	(1)	(2)	(3)
CERAC	1988 - 2002	2,127	2,127
CWT	1988 - 1991	1,694	1,132
COW	1988 - 1992	1,762	**2,444**
COW2	1988 - 1997	1,640	**2,214**
D&S	1988 - 1992	1,762	1,467
IISS	1988 - 2002	2,127	1,400
Ploughshares	1988 - 2002	2,127	1,282
SFTF	1988 - 2001	1,991	1,514 - 4,786
TPI	1988 - 1999	1,694	688
Uppsala/PRIO	1988 - 2002	2,127	545 - 1,000
UCDP	1989 - 2002	2,147	597 - 1,306
WMSE	1988 - 1990	1,638	**4,400**

(1) Overlap years between CERAC and the other datasets

(2) Annual average for CERAC during overlap years.

(3) Annual average for each dataset

Exploring the discrepancies

It is beyond the scope of this piece to attempt to reconcile CERAC numbers with all of the datasets considered above. Instead, we focus on UCDP. This choice is based on a number of considerations. First, the UCDP lower-bound curve resembles the CERAC figures, creating hope for reconciliation (Figure 3). Since both UCDP and CERAC work with events data, there are possibilities for detailed harmonization over the long run. Third, the UCDP concept of "battle deaths" resembles that of several other datasets examined above, including COW as well as Lacina and Gleditsch (2005) which we excluded because it already uses CERAC numbers. Therefore, the detailed comparison with UCDP should be representative of what might be obtained in a similar exercise for the other datasets. Fourth, the UCDP team was happy to work with us on this task.

It became immediately clear during our correspondence that the battle-death criterion is considerably more restrictive than CERAC's inclusion requirements. To a first approximation, battle deaths in the Colombian case cover only people killed, combatants or civilians, during direct combat events between government forces and left-

wing guerrillas. Upon closer examination the concept stretches a bit farther, for example, to include what CERAC considers one-sided events such as mine explosions and aerial bombardments. But UCDP excludes two broad categories of activity that CERAC includes. First, UCDP leaves out pure attacks on the civilian population as these do not involve battles. Second, UCDP excludes all activity of Colombia's illegal right-wing paramilitary groups, since they have no fundamental incompatibility with the state as UCDP requires.[12]

We applied the following procedures to pare the CERAC events list down to those that should survive the UCDP screen. We first classify all deaths in government-guerrilla clashes (battles), civilian or combatant, as battle related. There is just one subtlety that requires clarification. CERAC contains some compound events where there is an attack, by definition a one-sided event in CERAC, and a government-guerrilla clash in the same place and on the same day. For present purposes we treat such incidents as clashes and treat all deaths in the attack and clash components as battle related.

Next, we examined deaths in events that CERAC considers to be one-sided but which UCDP classifies as battle related, avoiding an arduous case-by-case determination. Rather, we considered first the type of each person killed in guerrilla attacks. Whenever any member of the government forces was killed in a guerrilla attack we treated all deaths in this event, including civilians, as battle-related. Otherwise, we excluded the event as an attack on civilians. For government attacks, of which a few aerial bombardments are the only significant ones, we followed a similar procedure. In events in which guerrillas were killed we treat all the deaths as battle-related. Otherwise we

[12] UCDP, in association with other institutions, is beginning to produce additional "human security" figures that are much closer to CERAC concepts but they are not yet publicly available.

excluded the event. The practical impact of this rule is that we dropped one government aerial bombardment that killed many paramilitaries but no guerrillas.

Figure 4 gives the modified CERAC time series together with the lower and upper estimates of UCDP figures. Modified CERAC is now contained within UCDP bounds or slightly above the UCDP upper bound in every year except 2002 when there is quite a substantial divergence. Moreover, modified CERAC dynamics look rather similar to lower-bound UCDP dynamics. The upsurge in violence from 1996 to 2002 is apparent in these curves but is much weaker than in the original CERAC curve. This is because much of the upsurge took the form or paramilitary attacks on civilians. The UCDP lower bound still remains substantially below the modified CERAC figure on average, as the harmonization of definitions closed only 56% of the gap between the two measures. Thus, the definitional differences are substantial but not the whole story. The remaining differences seem largely attributable to a fairly large number of small events that appear either in the local Colombian press or through CINEP networks, but that do not receive coverage in the international English-language press. In 2002 a really big gap opens up between CERAC and the UCDP curves. We would need to completely process the event list upon which the UCDP data is based to fully pin down the reason for this sudden disturbance. However, a plausible explanation is an overload of Colombia news in the international press during 2002 which was the hottest in the conflict by far and also included a pivotal presidential election, a series of local and national elections and a presidential inaugural ceremony marred by massive violence. Such an environment should push smaller conflict events out of the new.

Figure 4. Modified CERAC vs UCDP: new comparison

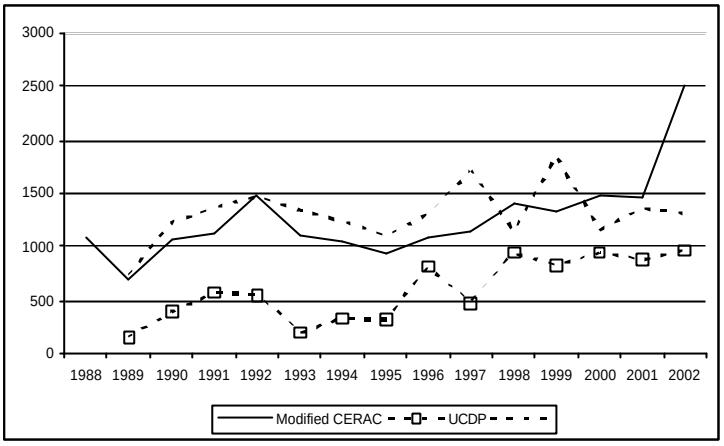


Table V follows Table IV except that the CERAC number is based on the modified CERAC series. Applying the strictness of the UCDP criteria eliminates circa 60% of the difference with UCDP. In fact, the procedure has taken the CERAC number slightly below four of the seven former underestimates so that the gap closed by 106% relative to these datasets. The only important overestimations are COW, COW2 and WMSE, the three datasets that were already above the regular CERAC numbers, with the gap increasing by 74% on average for these cases.

Table V. Annual averages for killing rates in overlap years: new comparison

Dataset	(1)	(2)	(3)
Modified CERAC*	1988 - 2002	1,264	1,264
CWT	1988 - 1991	995	**1,132**
COW	1988 - 1992	1,090	**2,444**
COW2	1988 - 1997	1,079	**2,214**
D&S	1988 - 1992	1,090	**1,467**
IISS	1988 - 2002	1,264	**1,400**
Ploughshares	1988 - 2002	1,264	**1,282**
SFTF	1988 - 2001	1,175	**1,514 - 4,786**
TPI	1988 - 1999	1,127	688
Uppsala/PRIO	1988 - 2002	1,264	545 - 1,000
UCDP	1989 - 2002	1,276	597 - 1,306
WMSE	1988 - 1990	953	**4,400**

(1) Overlap years between CERAC and the other datasets

(2) Annual average for CERAC during overlap years.

(3) Annual average for each dataset

*CERAC battle-deaths applying UCDP set of criteria

A Brief Northern Ireland Comparison

As a robustness check on our findings we look at the conflict in Northern Ireland, which has been extremely well documented by Malcolm Sutton (See Sutton, 1994 and http://cain.ulst.ac.uk/sutton/). Of the datasets covered above only UCDP and Uppsala/PRIO code Northern Ireland because the number of people killed is too small for inclusion according to the stricter criteria of other datasets.[13] Table VI gives time series for UCDP, Uppsala/PRIO, Sutton, and a modified Sutton concept applying similar procedures to those used to modify CERAC.[14] Consistent with the results for Colombia, the Sutton numbers are significantly higher than the UCDP figures but the gap closes by 80% on average with our modification procedures. An exception is 1998 where our modification procedures screen out a large event, the bombing of the Omagh shopping

[13] Lacina and Gleditsch (2005) do include Northern Ireland in their dataset, using precisely the Sutton figures.

[14] Specifically, we excluded events in which only civilians were killed and events involving Loyalist paramilitaries. Thus, we only included events where Republican paramilitaries killed British security force personnel and vice versa.

mall that killed 29 civilians and no security personnel, which UCDP leaves in. In fact, UCDP methodology specifically accepts terrorist incidents, something our modification procedures cannot detect.[15] Had we left this event in the gap would have closed by 89%.

Table VI. Total Annual deaths for Northern Ireland

year	Sutton	Uppsala/PRIO	UCDP	Gap Sutton-UCDP	Modified Sutton	Gap modified Sutton-UCDP	Gap reduction
1989	75	25 - 1,000	30	**45**	41	**11**	76%
1990	81	25 - 1,000	25	**56**	44	**19**	66%
1991	96	25 - 1,000	29	**67**	33	**4**	94%
1992	89	25 - 1,000	24	**65**	24	**0**	100%
1993	88	25 - 1,000	17	**71**	19	**2**	97%
1994	64		17	**47**	11	**6**	87%
1995	9		0	**9**	1	**1**	89%
1996	18		4	**14**	3	**1**	93%
1997	21		3	**18**	3	**0**	100%
1998	55	25 - 1,000	28	**27**	1	**27**	0%

We draw two main conclusions from the Northern Ireland case. First, Sutton's uses a broader conflict definition than UCDP's battle death concept. In this sense CERAC's approach to Colombia is similar to Sutton's approach to Northern Ireland and contrasts with UCDP's approach. Second, UCDP is measuring well what it intends to measure in the Northern Ireland case. This is not surprising, since the Northern Irish conflict is very well researched and the documentation is in English.

Conclusion

[15] Note that adding terrorist events, however definte, into the modified CERAC series would increase its distance from the UCDP series.

We have produced a snapshot of the world of cross-country conflict datasets. Our analysis suggests that these tend to underestimate the magnitude of the Colombian conflict and miss the significant upsurge in activity between 1996 and 2002. Our detailed comparison with UCDP data reveals that much, but not all, of the divergences are due to definitional differences.

But definitions are not simply matters of taste. Definitions must be appropriate for the questions researchers wish to answer. The UCDP battle death concept may be appropriate, for example, for evaluating the performance of government security forces in the field. But looking only at battle deaths can give a misleading picture of the overall situation in an irregular war. Armed groups often massacre civilians purposively to intimidate civilians into supporting their side in the battle (Kalyvas, 1999). Moreover, illegal right-wing paramilitary groups can form precisely due to limits on the extent to which government forces do participate in this dirty war (Mandler and Spagat, 2003). When such considerations operate, as in Colombia and a number of other irregular wars, a full quantitative approach should incorporate attacks on civilians and illegal paramilitary activity supporting the state side in a conflict. This is particularly important in the case of Colombia, where the right-wing paramilitaries are responsible for the biggest portion of civilian deaths (Restrepo and Spagat, 2004b).

Cross-country datasets have been instrumental in expanding our understanding of civil wars.[16] Econometric and statistical analyses of these datasets have generated much stimulating insight and debate (Collier and Hoeffler, 2004; F&L; Elbadawi and Sambanis, 2002). Most of these econometric studies of civil wars do not fall afoul of our critique, as they have simply used information on whether or not countries are at war at

[16] These insights are summarised in World Bank (2003) and Fitzgerald and Stewart (2000).

particular points in time. In fact, this focus is sensible given the limitations of the cross-country datasets highlighted in our paper. There are, however, exceptions such as Hartzell and Hoddie (2003) and Mesquida and Wiener (1999) that use conflict intensities.

We believe that significant further progress in civil war research will require improvements and extensions of existing datasets and the development of new ones so that investigators can open up the black box of conflict intensity and its dynamics. The key to this research programme is the construction of more micro datasets similar to CERAC. The development of new machine-coding technologies can support both new data construction and the quality checking of hand-coded datasets. There is ample room for many researchers and a variety of methodological approaches in this endeavour and we hope that many people will follow this path.

References

Bond, Doug, Joe Bond, Churl Oh, J. Craig Jenkins, & Charles Lewis Taylor, 2003. 'Integrated Data for Events Analysis (IDEA): An Event Typology for Automated Events Data Development', *Journal of Peace Research* 40 (6): 733-745.

Brunborg, Helge, Torklid H. Lyngstad & Henrik Urdal, 2003. "Accounting for genocide: How many were killed in Srebrenica?" *Journal of European Population* 19 (3): 229-248.

Buhaug, Halvard & Scott Gates, 2002. 'The Geography of Civil War'. *Journal of Peace Research* 39 (4): 417-433.

Collier, Paul & Anke Hoeffler, 2004. 'Greed and Grievance in Civil Wars', *Oxford Economic Papers* 56: 563-595.

Collier, Paul, Anke Hoeffler & Mäns Söderbom, 2004. 'On the Duration of Civil War', *Journal of Peace Research* 41 (3): 253-273.

Davenport, Christian, & Patrick Ball, 2002. 'Views to a Kill: Exploring the Implications of Source Selection in the Case of Guatemalan State Terror, 1977-1995', *Journal of Conflict Resolution* 46 (3): 427-450.

Doyle, Michael & Nicholas Sambanis, 2000. 'International Peacebuilding: A Theoretical and Quantitative Analysis', *American Political Science Review* 94 (4): 779-801.

Eck, Kristine, 2003. 'Conflict Dataset Catalogue', manuscript, Department of Peace and Conflict Research, Uppsala University (www.pcr.uu.se/research/UCDP/conflict_dataset_catalog/confliktdatasetcatalog.pdf)

Elbadawi, Ibrahim & Nicholas Sambanis, 2002. 'How Much War Will We See? Explaining the Prevalence of Civil War', *Journal of Conflict Resolution* 46(3): 307-334.

Eriksson, M. & P. Wallensteen, 2004. 'Armed Conflict, 1989–2003', *Journal of Peace Research* 41: 625-636.

Fearon, James D., 2004. 'Why Do Some Civil Wars Last so Much Longer Than Others?' *Journal of Peace Research* 41(3), 275–301.

Fearon, James D. & David Laitin, 2003. 'Ethnicity, Insurgency, and Civil War' *American Political Science Review* 97 (1): 75-90.

Fitzgerald, Edmund Valpy & Frances Stewart, 2000. *War and Underdevelopment, Volume 1. The Economic and Social Consequences of Conflict.* Oxford: Oxford University Press.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg & Håvard Strand, 2002. 'Armed Conflict 1946–2001: A New Dataset' *Journal of Peace Research* vol. 39 (5): 615-637.

Hartzell, Caroline, & Matthew Hoddie, 2003. 'Institutionalizing Peace: Power Sharing and Post-Civil War Conflict Management', *American Journal of Political Science* 47(2), 318-32.

IISS, 2004. *Armed Conflict Database.* Web-based dataset, The International Institute for Strategic studies (http://www.iiss.org/showpage.php?pageID=25).

Kalyvas, Stathis N., 1999 'Wanton and Senseless?: The Logic of Massacres in Algeria', *Rationality and Society* 11 (3): 243-286.

King, Gary & Will Lowe, 2003. 'An Automated Information Extraction Tool For International Conflict Data with Performance as Good as Human Coders: A Rare Events Evaluation Design', *International Organization* 57 (3): 617-642.

Lacina, Bethany & Nils Petter Gleditsch, 2005. 'Monitoring Trends in Global Combat: A New Dataset of Battle Deaths', *European Journal of Population*, forthcoming.

Licklider, Roy, 1995. 'The Consequences of Negotiated Settlements in Civil Wars, 1945-1993' *American Political Science Review* 89 (September): 681-690.

Mandler, Michael & Michael Spagat, 2003. 'Foreign Aid Designed to Diminish Terrorist Atrocities can Increase them', CEPR Discussion Paper 4004.

Marshall, Monty G., Ted Robert Gurr &and Barbara Harff, 2002. 'State Failure Task Force Problem Set Codebook: Internal Wars and Failures of Governance, 1955-2002', manuscript (http://www.cidcm.umd.edu/inscr/stfail/sfcodebk.htm).

Mesquida, Christian G., & Neil I. Weiner, 1999. 'Male Age Composition and Severity of Conflicts', *Politics and the Life Sciences* 18(2), 113-7.

Mueller, John, 1995. 'The Perfect Enemy: Assessing the Gulf War', *Security Studies* 5(1): 79-117.

Posada, Eduardo, 2001. 'Civil War? The Language of Conflict in Colombia', manuscript, Fundación Ideas para la Paz, Bogotá. Colombia.

Project Ploughshares, 2004. *Armed Conflicts Report 2003*. Waterloo, Canada: Project Ploughshares, Institute of Peace and Conflict Studies.

Rabasa, Angel & Peter Chalk, 2001. 'Colombian Labyrinth', Santa Monica, CA: RAND.

Regan, Patrick, 2002. 'Third Party Interventions and the Duration of Civil Wars', *Journal of Conflict Resolution* 46 (1): 55-73.

Restrepo, Jorge, Michael Spagat & Juan F. Vargas, 2004. 'The Dynamics of the Colombian Civil Conflict: A New Dataset', *Homo Oeconomicus* 21 (2): 396-428.

Restrepo, Jorge & Michael Spagat, 2004a. 'The Colombian Conflict: Uribe's First 17 Months', manuscript, Royal Holloway College, University of London.

Restrepo, Jorge & Michael Spagat, 2004b. 'Civilian Casualties in the Colombian Conflict: A New Approach to Human Security', manuscript, Royal Holloway College, University of London.

Riascos, Alvaro & Juan F. Vargas, 2004. 'Violence and Growth in Colombia: A Brief Review of the Literature', *Webpondo* (www.webpondo.org), January-March.

Roberts, Les, Pascal Ngoy, Colleen Mone, Charles Lubula, Luc Mwezse, Mariana Zantop, & Michael Despines, 2003. *Mortality in the Democratic Republic of Congo: Results from a Nationwide Survey.* International Rescue Committee, Bukavu/New York.

Roberts, Les, Lafta Riyadh, Richar Garfeld, Jamal Khudhairi & Gilbert Burnham, 2004. *Mortality Before and After the 2003 Invasion of Iraq: Cluster Sample Survey*. Center for

International Emergency Disaster and Refugee Studies. Johns Hopkins Bloomberg Scholl of Public Health.

Sarkees, Meredith. R., 2000. 'The Correlates of War Data on War: An Update to 1997', *Conflict Management and Peace Science* 18 (1): 123-144.

Schrodt, Philip A. & Deborah Gerner, 1994. 'Validity Assessment of a Machine-Coded Event Data Set for the Middle East, 1982–92', *American Journal of Political Science* 38(3): 825–854.

Sivard, Ruth, 1991. *World Military and Social Expenditures.* Washington: World Resources Institute.

Small, Melvin & J. David Singer, 1982. *Resort to Arms: International and Civil Wars, 1816-1980*. Beverly Hills, CA: SAGE.

Sutton, Michael, 1994. *Bear in mind these dead ... An Index of Deaths from the Conflict in Ireland 1969-1993*. Belfast: Beyond the Pale.

UCDP (2004). *The Uppsala Conflict Database.* Web-based dataset, Uppsala: The Uppsala Conflict Data Project, Department of Peace and Conflict Research, Uppsala University (http://www.pcr.uu.se/database/index.php).

World Bank, 2003. *Breaking the Conflict Trap: Civil War and Development Policy.* Washington: World Bank Policy Report.