Web Appendix "Accounting for the Dynamics of One-Sided Violence: Introducing KOSVED" (Gerald Schneider/Margit Bussmann)

This online appendix consists of several documents which are also available from the Web site of the KOSVED project.

A) Codebook
B) Coders and sources
C) Comparison with other data sets

The Konstanz One-Sided Event Dataset (KOSVED) is the result of a data gathering effort that was conducted at the Department of Politics and Management at the University of Konstanz from October 1, 2007, to March 31, 2010. The project team consisted of Margit Bussmann (now University of Greifswald) and Gerald Schneider as principal investigators and several assistants and coders.

KOSVED provides detailed information on the magnitude and location of events of one-sided violence in 17 civil wars. Schneider and Bussmann (2012) describe the data set in detail. Information on the variables can be found in the KOSVED codebook.

When using KOSVED, please cite the data set article: Schneider, Gerald and Margit Bussmann. 2012. "Accounting for the Dynamics of One-Sided Violence: Introducing KOSVED". Unpublished manuscript, University of Konstanz/University of Greifswald and, if needed, the codebook: Margit Bussmann and Gerald Schneider (with Alexander Bräunig, Constantin Ruhe, Adam Scharpf and Roos van der Haer). 2011. Konstanz One-Sided Violence Event Dataset (KOSVED) Codebook. Version 2.0

Appendix A: Konstanz One-Sided Violence Event Dataset (KOSVED) Codebook
Version 1.2 – January 27, 2012

Margit Bussmann
University of Greifswald,
Department of Politics and Communication

&

Gerald Schneider
University of Konstanz,
Department of Politics and Management

(with Alexander Bräunig, Constantin Ruhe, Adam Scharpf and Roos van der Haer)

1. Introduction

This document describes the Konstanz One-Sided Violence Event Dataset (KOSVED), a project financed by the German Peace Foundation (DSF) and conducted at the Department of Politics and Management, University of Konstanz, in the period from October 1, 2007, to March 31, 2010. The project team consisted of Margit Bussmann and Gerald Schneider as principal investigators and several student assistants who are listed on the project homepage (http://www.polver.uni-konstanz.de/gschneider/kosved). While this document describes KOSVED in detail, Schneider and Bussmann (2012) offer insight in the data set, comparison with related data gathering efforts and discussion of the reliability of the measures introduced below.

The unit of analysis in this KOSVED dataset is the individual event of one-sided violence. The dataset is constructed so as to be compatible with the Armed Conflict Location and Events Dataset (ACLED), which is mainly focused on individual battle events, transfer of military control from the government to the rebel groups and vice versa, and the location of rebel group strongholds (see Raleigh et al. 2010). KOSVED enlarges this dataset by making more detailed information on one-sided violence events available. It describes not only the location and time of such events but also provides as detailed as possible information on the type of weapons used in the conduct of the atrocities, the number of civilians killed or harmed as well as the nationality, ethnicity, gender and age of the civilian targets. If possible, we have also collected the nationality, ethnicity, and the formal organizational name of the perpetrators and whether these actors were affiliated with the government or the rebels in the conflict under scrutiny. KOSVED also offers information on tactics that accompanied the one-sided violence; such as for example looting and kidnapping. Finally, we provide information on the number of news reports that mentioned the event, whether the status of the event is contested, and whether the report is based on a primary or secondary account.

2. Terms and Definitions

The project's definition of one-sided violence builds on Eck and Hultman (2007). Schneider and Bussmann (2012) discuss in detail where the KOSVED definition of one-sided violence differs from the one of Eck and Hultman (2007). The most important addition to their definition of one-sided violence is the distinction between violence directed at civilians either as primary or secondary targets.

One-sided violence
Violent acts perpetrated by an organized group, which can be either a rebel organization or government troops, directed against a group of unarmed non-combatants during, shortly before, or after a conflict. These acts result in the immediate physical harming or death of more than one non-combatant.

The separate elements of the definition are operationalized as follows:

- *Violent acts*: The usage of any form of physical force to inflict immediate severe harm or death to a non-combatant. The arms used to this end include any material means, e.g. manufactured weapons but also sticks, stones etc. employed to harm or kill civilians or unarmed soldiers.

- *Organized group*
    - o *Rebel*: A group of people who employ a distinctive name and who are challenging the authority of the government or a competing group through the threat to use

violence or through the usage of it. This definition includes militias operating in conjunction or alliance with the rebel group. These groups are supported, armed, or allied to the rebel forces.

- o *Government*: An internationally recognised regime that is assumed to control the territory of the state and whose sovereignty is not disputed by another internationally recognised regime. The groups included in this definition are typically official state actors, like the military or police forces, but can also be militias operating in conjunction or alliance with the recognized government. These groups are supported, armed, or allied to the government forces.

- *Non-combatant*: Following Valentino (2004: 8), a non-combatant is defined as "any unarmed individual who is not a member of a professional or guerrilla military group and who does not actively participate in hostilities by tending to cause physical harm to enemy personnel or property."

- *Conflict*: As defined by the UCDP/PRIO data set, a conflict is any incompatibility concerning government and territory contested by arms with at least 25 battle-related deaths in a given year. This version of KOSVED includes internal armed conflicts and internationalised internal armed conflicts (i.e. internal conflicts with intervention from other states) but excludes extra-systemic conflicts and interstate conflicts (see Gleditsch et al. 2002). To make sure that the process leading to the outburst of conflict and to the stabilisation of peace is sufficiently surveyed we coded two years before the onset of the (internationalised) internal armed conflict and two years after its conclusion according to the length of conflict reported in the UCDP/PRIO data set (Harbom & Sundberg, 2009/ Version 4-2007). Most conflicts were selected on the basis of their severity but others were selected based on their reputation of cruelty against civilians, i.e. some minor conflicts were selected since we expected a considerable amount of one-sided violence.

3. Variables in the Konstanz One-Sided Violence Event Dataset

Table 1: Definition of Variables in the KOSVED dataset

Variable	Label	Brief description
Country	Country	The name of the country in which the event took place.
CountryPart	CountryPart_	The part of the country (North, South, East, or West) where the event took place.
Region	Region	The name of the region in which the event took place.
RegionPart	RegionPart_	The part of the region (North, South, East, or West) where the event took place.
Place	Place	The name of the place in which the event took place.
Date	Date	The date of the event.
Precision	Precision_	The level of precision for the date of the event. Possible precision levels: year, month, week or day.
Week	Week	Part of the month in which the event took place.
ReportBase	ReportBase	Whether the event was directly witnessed or not.
Status	Status	Whether the event was contested by one of the warring parties or not.
NumberReports	NumberReports	The number of reports that mention the event.
TargetReport	TargetReport	Civilians targeted as primary or secondary target.
CivilDefintion	CivilianDefinition	Description of target as "civilian" or as "people" or not further detailed.
DPrimitive	D_Primitive	The number of civilians killed by primitive weapons.

DShooting	D-Shooting	The number of civilians killed by shooting.
DBombing	D-Bombing	The number of civilians killed by bombings.
DSuicide	D-Suicide	The number of civilians killed by suicide attacks.
DTank	D-Tank	The number of civilians killed by tank attacks.
DAircraft	D-Aircraft	The number of civilians killed by aircraft attacks.
DWMD	D-WMD	The number of civilians killed by weapons of mass destruction.
DUnclear	D-Unclear	The number of civilians killed by unclear or unknown means.
HSexual	H-Sexual	The number of civilians harmed by sexual violence.
HPrimitive	H-Primitive	The number of civilians harmed by primitive weapons.
HShooting	H-Shooting	The number of civilians harmed by shooting.
HBombing	H-Bombing	The number of civilians harmed by bombings.
HSuicide	H-Suicide	The number of civilians harmed by suicide attacks.
HTank	H-Tank	The number of civilians harmed by tank attacks.
HAircraft	H-Aircraft	The number of civilians harmed by aircraft attacks.
HWMD	H-WMD	The number of civilians harmed by weapons of mass destruction.
HUnclear	H-Unclear	The number of civilians harmed by unclear or unknown means.
TEthnicity	T-Ethnicity	The ethnicity of the targets.
TNationality	T-Nationality	The nationality of the targets.
TMaleGeneral	T-MaleGeneral_	The number of males as targets.
TMaleChild	T-MaleChild_	The number of boys as targets.
TMaleAdult	T-MaleAdult_	The number male adults as targets.
TMaleElders	T-MaleElders_	The number of male elderly as targets.
TFemaleGeneral	T-FemaleGeneral_	The number of females as targets.
TFemaleChild	T-FemaleChild_	The number of girls as targets.
TFemaleAdult	T-FemaleAdult_	The number of female adults as targets
TFemaleElders	T-FemaleElders_	The number of female elders as targets.
TUnclearGeneral	T-UnclearGeneral_	The number of persons targeted of unknown gender.
TUnclearChild	T-UnclearChild_	The number of children with unknown gender.
TUnclearAdult	T-UnclearAdult_	The number of adults with unknown gender.
TUnclearElders	T-UnclearElders_	The number of elderly with unknown gender.
TCondition	T-Condition	The condition of targets before the event.
PName	P-Name	The formal name of the perpetrator of the event.
PAcronym	P-Acronym	The acronym of the perpetrator of the event.
PEthnicity	P-Ethnicity	The ethnicity of the perpetrator of the event.
PNationality	P-Nationality	The nationality of the perpetrator of the event.
PStateChar	P-StateChar_	The state characteristics of the perpetrator
PNonstateChar	P-NonstateChar_	The non-state characteristics of the perpetrator.
PNonstatChar	P-NonstateChar_	Non-state actor code of UCDP Non-state Actor Dataset v. 1-2009
Govobserv	Govobserv	The presence of governmental observers during an event.
NGOobserv	NGOobserv	The presence of NGOs during an event.
IGOobserv	IGOobserv	The presence of IGOs during an event
Profitobserv	Profitobserv	The presence of profit organizations during an event.
RLooth	R-Looth	The use of looting as a related tactic.
RScorched	R-Scorched	The use of scorched earth tactics as a related strategy.
RStarvation	R-Starvation	The use of starvation as a related tactic.
RShields	R-Shields	The use of human shield as a related tactic.
RAssasin	R-Assasin	The use of assassination attempts as a related tactic.
RTorture	R-Torture	The use of torture as related tactic.

RHostage	R-Hostage	The use of hostage taking as a related tactic.
RKidnapping	R-Kidnapping	The use of kidnapping as a related tactic.
RDeport	R-Deport	The use of deportation and forced migration as a related tactic.
RDisappear	R-Disappear	The disappearance of non-combatants.
RCamps	R_Camps	The use of camps as a related tactic.
RSieges	R_Sieges	The use of sieges as a related tactic.

3.1 Geographical information

- *Country*: The geographical country where the one-sided violence event took place. Country is a string variable.
- *CountryPart*: In some news reports it is mentioned that the one-sided violence event took place in a specific part of the country. For the cases of Bosnia and Croatia these information are missing. Four dummy variables where created to capture this information:
    - *CountryPart_North*: This variable is coded as "North" if the one-sided violence event took place in the North of the country, . otherwise.
    - *CountryPart_South*: This variable is coded as "South" if the one-sided violence event took place in the South of the country, . otherwise.
    - *CountryPart_East:* This variable is coded as "East" if the one-sided violence event took place in the East of the country, . otherwise.
    - *CountryPart_West:* This variable is coded as "West" if the one-sided violence event took place in the West of the country, . otherwise.
- *Region*: The name of the geographical region in which the one-sided violence event took place. Region is a string variable.
- *RegionPart:* Some news reports mention that the one-sided violence event took place in a specific part of the region. For the cases of Bosnia and Croatia these information are missing. Four dummy variables were created to capture this information:
    - *RegionPart_North*: This variable is coded as "North" if the one-sided violence event took place in the North of the region, . otherwise.
    - *RegionPart_South*: This variable is coded as "South" if the one-sided violence event took place in the South of the region, . otherwise.
    - *RegionPart_East*: This variable is coded as "East" if the one-sided violence event took place in the East of the region, . otherwise.
    - *RegionPart_West*: This variable is coded as "West" if the one-sided violence event took place in the West of the region, . otherwise
- *Place*: The name of the geographical place in which the one-sided violence event took place. Place is a string variable.

3.2 Time information

- *Date*: The date, in a 'yyyymmdd format', of the one-sided violence event.
    - *Precision*: The level of precision for the date of the event. For some events, it is impossible to pinpoint the exact day, week, or even month due to a lack of precise information. Precision is coded to highlight the level of certainty for the date variable. It is a string variable that indicates whether the precision is at the level of "year", "month", "week", or "day".
- *Week*: Some reports indicate the time frame during which the event allegedly took place. One example is the statement: "in the beginning of January". These instances are coded as happening in the <u>first week</u> (coded as 1). Reports about one-sided events in "mid-February" will be coded as the event happening in the <u>second or third</u> week of the month

(coded as 2). Lastly, sometimes it is mentioned that the one-sided violence event took place in "late-March". These instances are coded for the <u>last week</u> of that particular month (coded as 3).

Note that we randomly selected dates when some information on a specific event data was missing for the temporal aggregation of the data set. The country-specific information sheets available on the project website inform about the details of these data imputations and their frequency.

3.3   Reporting

- *Reportbase:* This indicator looks at the empirical basis of the reports. Some news reports for instance quote a witness who described the event or in which the journalist him/herself witnessed the act of one-sided violence (coded as 1). Other reports only mention the one-sided violence event on the basis of a secondary account (coded as 2). *Status:* If one of the warring parties denies the existence or responsibility of an act of one-sided violence, we code this as a contested event (coded as 2). If the event was not contested by any of the parties involved it was coded as 1. Note that this dummy variable does not capture possible disagreements about the number of non-combatants killed or harmed.
- *NumberReports:* The number of news reports that mention the one-sided violence event.
- *TargetReport:* This category distinguishes whether the event was an instance of one-sided violence with the killing and harming of non-combatants as primary or secondary target. If it was clear from the news reports that the perpetrator had the intention to kill or harm non-combatants it was coded as 1 (primary target). These cases of one-sided violence were only entered in the dataset as long as they resulted in the killing/harming of more than one civilian. This allowed us to exclude assassination attempts. Sometimes, the killing or harming of innocent civilians was collateral damage resulting from a violent action directed against another primary aim such as a military unit (coded as 2). Then the civilian was considered "secondary target". We entered these events as long as they resulted in more than nine civilians injured or killed. This number is of course arbitrary, but close to the threshold used in other datasets. When we could not establish whether the civilians were either primary or secondary targets, this variable received a value of 3.

CivilDefintion: This variable captures how the news reports describe the targets of acts of one-sided violence. Sometimes these targets are described as "civilians" or "non-combatants" (coded as 1) or as "people", "farmers" etc. (coded as 2). This distinction is important as it is often impossible to extract whether "people" are part of the warring parties or not. Sometimes, however, this distinction was difficult to make. This is for instance the case when the news report under consideration talks about "several bodies lying on the street". In these instances, the civilian status was difficult to determine, and therefore we coded them as 3.

3.4   One-sided violence resulting in the death of non-combatants

The following variables specify the number of non-combatants killed in general and with a specific weapon. In case we have several news reports on the same event with differing numbers of civilians harmed and killed, we list the minimum, maximum, average and modus number of victims given in the reports. All these killing-variables have five categories:

- ➢ D-(Weapon)Min: The minimum number of non-combatants killed.
- ➢ D-(Weapon)Max: The maximum number of non-combatants killed.
- ➢ D-(Weapon)Av: The average number of non-combatants killed.
- ➢ D-(Weapon)Mod: The modus number of non-combatants killed.
- ➢ D-(Weapon)Index: Some reports do not include a clear number of killed non-combatants, but provide a verbal estimate instead. When news reports mention keywords like "few", "a number of", "several", the event is coded as 1. Keywords like

"ten" or "a dozen" lead to a coding in the category "2", while keywords like "tens" and "scores" are coded as 3. We classify events which are accompanied by descriptions like "hundreds" and "tens of dozens" into the category "4". "Thousands" are coded as 5, and words like "tens of thousands" are coded as 6.

The number of non-combatants killed according to type of weapons

- *D-Primitive:* The number of non-combatants killed by the use of primitive weapons such as spears, axes, machetes, and clubs. (*D-PrimitiveMin/ D-PrimitiveMax/ D-PrimitiveAv/ D-PrimitiveMod/ D-PrimitiveIndex*).
- *D-Shooting*: The number of non-combatants killed through firearms (e.g., pistols, rifles, automatic firearms). (*D-ShootingMin/ D-ShootingMax/ D-ShootingAv/ D-ShootingMod/ D-ShootingIndex*).
- *D-Suicide*: The number of non-combatants killed by the use of explosive devices that also take the lives of the perpetrators. (*D-SuicideMin/ D-SuicideMax/D-SuicideAv/ D-SuicideMod/D-SuicideIndex*).
- *D-Bombing*: The number of non-combatants killed by means of bombs, grenades, or other explosive devices. The event occurred without the suicide(s) of the perpetrator(s). (*D-BombingMin/D-BombingMax/D-BombingAv/D-BombingMod/D-BombinIndex*).
- *D-Tank*: The number of non-combatants killed by means of tank attacks. (*D-TankMin/ D-TankMax/ D-TankAv/ D-TankMod/ D-TankIndex*).
- *D-Aircraft*: The number of non-combatants killed by means of aircraft attacks. (*D-AircraftMin/D-AircraftMax/ D-AircraftAv/D-AircraftMod/D-AircraftIndex*).
- *D-WMD:* The number of non-combatants killed by weapons of mass destruction (chemical, biological or radiological weapons). (*D-WMDMin/ D-WMDMax/ D-WMDAv/D-WMDMod/D-WMDIndex*).
- *D-Unclear*: The number of non-combatants killed by unclear, unknown or other means. (*D-UnclearMin/D-UnclearMax/D-UnclearAv/D-UnclearMod/D-UnclearIndex*).

3.5 One-sided violence resulting in injured non-combatants

Like for deadly events of one-sided violence, the following variables list the number of harmed non-combatants per weapon mentioned in the news reports. All these variables have five categories:

- ➤ H-(Weapon)Min: The minimum number of non-combatants harmed.
- ➤ H-(Weapon)Max: The maximum number of non-combatants harmed.
- ➤ H-(Weapon)Av: The average number of non-combatants harmed.
- ➤ H-(Weapon)Mod: The modus number of non-combatants harmed.
- ➤ H-(Weapon)Index: Some reports do not include a clear number of harmed non-combatants, but provide a verbal estimate instead. When news reports mention keywords like "few", "a number of", "several", the event is coded as 1. Keywords like "ten" or "a dozen" lead to a coding in the category "2", while keywords like "tens" and "scores" are coded as 3. We classify events which are accompanied by descriptions like "hundreds" and "tens of dozens" into the category "4". "Thousands" are coded as 5, and words like "tens of thousands" are coded as 6.

The number of non-combatants harmed according to type of weapons

- *H-Sexual*: The number of non-combatants harmed by sexual violence (harm by raping, coerced undressing and non-penetrating sexual assault such as sexual mutilation (Wood 2006)). (*H-SexualMin /D-SexualMax /H-SexualAv /H-SexualMod/H-SexualIndex*).

- *H-Primitive:* The number of non-combatants harmed by the use of primitive weapons such as spears, axes, machetes, and clubs. (*H-PrimitiveMin /D-PrimitiveMax /H-PrimitiveAv /H-PrimitiveMod/H-PrimitiveIndex).*

- *H-Shooting*: The number of non-combatants harmed through firearms (e.g., pistols, rifles, automatic firearms). (*H-ShootingMin/H-ShootingMax/H-ShootingAv/H-ShootingMod/ H-ShootingIndex).*

- *H-Suicide*: The number of non-combatants harmed by the use of explosive devices that also take the lives of the perpetrators. (*H-SuicideMin/ H-SuicideMax/H-SuicideAv/ H-SuicideMod/ H-SuicideIndex).*

- *H-Bombing*: The number of non-combatants harmed by means of bombs, grenades, or other explosive devices. This is done without the suicide(s) of the perpetrator(s). (*H-BombingMin/ H-BombingMax/ H-BombingAv/ H-BombingMod/H-BombinIndex)*

- *H-Tank*: The number of non-combatants harmed by means of tank attacks. (*H-TankMin/ H-TankMax/ H-TankAv/H-TankMod/H-TankIndex).*

- *H-Aircraft*: The number of non-combatants harmed by means of aircraft attacks. (*H-AircraftMin/ H-AircraftMax/ H-AircraftAv/ H-AircraftMod/ H-AircraftIndex).*

- *H-WMD:* The number of non-combatants harmed by weapons of mass destruction (chemical, biological or radiological weapons). (*H-WMDMin/ H-WMDMax/ H-WMDAv/ H-WMDMod/ H-WMDIndex)*

- *H-Unclear*: The number of non-combatants harmed by unclear, unknown or other means. (*H-UnclearMin/ H-UnclearMax/ H-UnclearAv/ H-UnclearMod/ H-UnclearIndex).*

3.6   Target information

- *T-ethnicity*: The ethnicity of the targeted civilians. T-ethnicity is a string variable.
- *T-nationality*: The nationality of the targeted civilians. T-nationality is a string variable.

The following variables disaggregate the number of killed and harmed non-combatants by gender and age group if such information appears in the news reports. All these variables have five categories: they report the minimum, maximum, average, modus, and/or a possible index of civilian casualties (harming and killing).

- *T-MaleGeneral*: The number of killed or harmed male non-combatants. (*T-MaleGeneralMin/ T-MaleGeneralMax/ T-MaleGeneralMod/ T-MaleGeneralAv/ T-MaleGeneralIndex).*

- *T-MaleChild:* The number of killed or harmed boys (age 0-16). (*T-MaleChildMin/ T-MaleChildMax/ T-MaleChildMod/ T-MaleChildAv/T-MaleChildIndex).*

- *T-MaleAdult:* The number of killed or harmed male adults (age 17-54). (*T-MaleAdultMin/ T-MaleAdultMax/ T-MaleAdultMod/ T-MaleAdultAv/T-MaleAdultIndex).*

- *T-MaleElders:* The number of killed or harmed male elderly (age 55+). (*T-MaleEldersMin/ T-MaleEldersMax/ T-MaleEldersMod/ T-MaleEldersAv/T-MaleEldersIndex).*

- *T-FemaleGeneral*: The number of killed or harmed female non-combatants. (*T-FemaleGeneralMin/ T-FemaleGeneralMax/ T-FemaleGeneralMod/ T-FemaleGeneralAve/ T-FemaleGeneralIndex).*

- *T-FemaleChild:* The number of killed or harmed girls (age 0-16). (*T-FemaleChildMin/ T-FemaleChildMax/ T-FemaleChildMod/ T-FemaleChildAv/T-FemaleChildIndex).*

- *T-FemaleAdult:* The number of killed or harmed female adults (age 17-54). (*T-FemaleAdultMin/ T-FemaleAdultMax/ T-FemaleAdultMod/ T-FemaleAdultAv/ T-FemaleAdultIndex).*

- *T-MaleElders:* The number of killed or female elderly (age 55+). (*T-FemaleEldersMin/ T-FemaleEldersMax/ T-FemaleEldersMod/ T-FemaleEldersAv/T-FemaleEldersIndex).*

- *T-UnclearGeneral*: The number of killed and harmed civilians with unknown gender.

(T-UnclearGeneralMin/ T-UnclearGeneralMax/ T-UnclearGenderAv/T-UnclearGenderMod/ T-UnclearGeneralIndex).

- *T-UnclearChild*: The number of killed and harmed children (age 0-16) with unknown gender. *(T-UnclearChildMin/ T-UnclearChildMax/ T-UnclearChildAv/T-UnclearChildMod/ T-UnclearChildIndex).*
- *T-UnclearAdult*: The number of killed and harmed adults (age 17-54) with unknown gender. *(T-UnclearAdultMin/ T-UnclearAdultMax/ T-UnclearAdultAv/T-UnclearAdultMod/ T-UnclearAdultIndex).*
- *T-UnclearElders:* The number of killed and harmed elderly (age 55+). *(T-UnclearEldersMin/ T-UnclearEldersMax/ T-UnclearEldersAv/T-UnclearEldersMod/ T-UnclearElderstIndex).*

- In the post-coding process we corrected the gravest of coding mistakes in the variables of Death and Harmed with a Consistency Test: In case the added number of the Death and Harmed was smaller than the number of Targets (dht-consistency ==-1), we set the number of targets down to the minimum of death plus harmed and sorted it into T-UnclearGeneral. This leads to a loss of information, but creates a consistent result for each event.

- *T-Condition*: The condition of the target(s) before the attack. This variable takes the value of 1 for previously injured civilians, 2 for ill civilians, 3 for pregnant women, 4 for prisoners of war and 5 when the condition of the targets is unknown or unclear.

3.7 Perpetrator information

- *P-Name*: The formal name of the organization that perpetrated the act of one-sided violence. This can be the name of a rebel movement but also the government of a particular country. P-Name is a string variable.
- *P-Acronym*: The acronym of the organization (state or non-state actor) that perpetrated the act of one-sided violence.
- *P-Ethnicity*: The ethnicity of the perpetrator. P-Ethnicity is a string variable
- *P-Nationality*: The nationality of the perpetrator. P-Nationality is a string variable.
- *P-StateChar*: Indicates whether the perpetrator of the act of one-sided violence was identified as a state actor.
  - *P-StateChar_Military*: Dummy variable coded as 1 if the attackers are identified as uniformed members of an official state military force, 0 otherwise.
  - *P-StateChar_Police*: Dummy variable coded as 1 if the attackers are identified as law enforcement officers, 0 otherwise.
  - *P-StateChar_Mercenaries*: Dummy variable coded as 1 if the attackers are identified as part of a private army or were mercenaries hired by the state, and 0 otherwise. We follow the definition of a mercenary according to the Additional Protocol I, Article 47, of the Geneva Conventions of 1949 (ICRC 1949/1977).
  - *P-StateChar_StateAgencies*: Dummy variable coded as 1 if the attackers are identified as employees of a state agency other than the military and police forces, 0 otherwise.
  - *P-StateChar_Unclear*: Dummy variable coded as 1 if the exact state agency of the attacker is unknown or unclear, 0 otherwise.
- *P-NonstateChar:* Indicates that the perpetrator of the one-sided violence event belongs to a non-state agency.

- o *P-NonstateChar_Top:* A dummy variable that is coded as 1 if attackers are identified as belonging to the top level of the rebel movement, 0 otherwise.
- o *P-NonstateChar_Medium:* A dummy variable that is coded as 1 if the attackers are identified as belonging to the medium level of the rebel movement, 0 otherwise.
- o *P-NonstateChar_Low:* A dummy variable that is coded as 1 if the attackers are identified as foot-soldiers of the rebel movement, 0 otherwise.
- o *P-NonstateChar_Mercenaries:* A dummy variable that is coded as 1 if the attackers are identified as hired mercenaries working for a non-state actor (see definition above), 0 otherwise.
- o *P-NonstateChar_Unclear:* A dummy variable that is coded as 1 if the attackers were part of the non-state armed movement but their exact position in the movement was unclear or unknown, 0 otherwise.

3.8 International Monitoring

We also included information on the possible presence of international actors at the time of the violent acts. While we did not systematically search with specific keywords, we incorporated these variables if our coders came across relevant information on international presence in one of the news reports.

- *Govobserv:* The name of the country of third party governmental observers who were present at the time of the event of one-sided violence. Presence means in close proximity (max. 100 km) to the place where the event happened. Govobserv is a string variable.
- *NGOobserv:* The name of the Non-Governmental Organization (NGO) which was present at the time of the event of one-sided violence. Presence means in close proximity (max. 100 km) to the place where the event happened. NGOobserv is a string variable that lists the formal names of all the NGOs present during the event.
- *IGOobserv:* The name of the Intergovernmental Organization (IGO) which was present at the time of the event of one-sided violence. Presence means in close proximity (max. 100 km) to the place where the event happened. IGOobserv is a string variable that lists the formal names of all the IGOs that were present during the event.
- *Profitobserv:* The name of the profit organization which was present during an event of one-sided violence. Presence means in close proximity (max. 100 km) to the place where the event happened. Profitobserv is a string variable that lists the formal names of all profit organizations that were present during the event.

3.9 Related tactics

In some instances, the perpetrators use tactics that accompany the use of violence against civilians. These related tactics are captured in dummy variables (1 if employed, 0 otherwise); they are only coded when they are mentioned in the context of an instance of one-sided violence.

- *R-Looth:* Looting means indiscriminate stealing of goods by force.
- *R-Scorched:* Scorched earth tactics include wide-scale burning of houses and other buildings, destruction of infrastructure such as roads, power lines, and water treatment systems, and other tactics used to make an area uninhabitable.
- *R-Starvation:* Tactics related to starvation include the deliberate withholding of food aid, deliberate destruction of crops, attacks on food-aid convoys, and interference with the delivery of medical or humanitarian aid.
- *R-Shields:* Using human shields involves moving non-combatants onto battlefield as protection of the armed group, hiding in hospitals, hiding among crowds, etc.
- *R-Assasin:* Assassinations are discrete events of killing a leader of one of the involved organized groups.

- *R-Torture*: Torture is a deliberate physical act to inflict severe suffering or pain on the human body. .
- *R-Hostage*: Taking hostages is the forceful abduction (of children or adults) until specific demands are met or money is handed over.
- *R-Kidnapping:* Kidnapping is the forceful abduction (of children or adults) for the use as fighters.
- *R-Deport:* Deportation is forced banishment of civilians from a certain area.
- *R-Disappear:* Non-combatants disappear without evidence of death or abduction. This will also occur when there are reports of additional victims who cannot be accounted for among the known dead.
- *R-Camps:* A guarded enclosure for the detention or imprisonment of political prisoners, prisoners or war, aliens, refugees.
- *R-Sieges:* Situations where the freedom of movement of non-combatants is restricted to a limited geographical area, but where the targeted individuals and groups were not forcefully transported to these locations. Usually the inhabitants maintain internal control of the area to which they are restricted because of the siege.

4. References

Eck, Kristine and Lisa Hultman 2007. One-Sided Violence against Civilians in War: Insights from new Fatality Data. *Journal of Peace Research* 44(2): 233-246.

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): 615-637.

Harbom, Lotta and Ralph Sundberg. 2011. UCDP Actor Dataset Codebook. Uppsala Conflict Data Program (UCDP)/Department of Peace and Conflict Research, Uppsala University (http://www.pcr.uu.se/digitalAssets/63/63658_UCDP_Actor_Dataset_Codebook_2011.pdf, last consulted January 27, 2012).

International Committee of the Red Cross 1949/1977. Protocol Additional to the Geneva Conventions of 12 August 1949, and relating to the Protection of Victims of International Armed Conflicts (Protocol I), 8 June 1977. (http://www.icrc.org/ihl.nsf/FULL/470?OpenDocument, last consulted January 27, 2012)

Raleigh, Clionadh, Andrew Linke, Håvard Hegre & Joachim Carlsen (2010) Introducing ACLED: An Armed Conflict Location and Event Dataset. *Journal of Peace Research* 47(5): 651-660.

Schneider, Gerald and Margit Bussmann. 2012. Accounting for the Dynamics of One-Sided Violence: Introducing KOSVED. Unpublished manuscript, University of Konstanz/University of Greifswald.

Valentino, Benjamin A. 2004. *Final Solutions: Mass Killings and Genocide in the Twentieth Century.* Ithaca, NY: Cornell University Press.

Wood, Elisabeth Jean. 2006. Variation in Sexual Violence during War. *Politics & Society* 34(3): 307-341.

Appendix B: KOSVED (Version 1.2, February 2012) sources and coders
Student assistants working for the Konstanz One-Sided Violence Event Data Set (KOSVED) coded with the help of several media instances of violence against civilians and unarmed combatants. Below you will find a detailed list of the sources used to code events of violence against civilians in a particular conflict. The coders relied for the analysis of all conflicts on *BBC Monitoring* and *The New York Times*. For African conflicts, we added *Africa News* to these sources and *Associated Press Worldstream* for the other civil wars covered. Some of the coders were able to research additional sources. The list below details the sources used per conflict and in parentheses the coders. Note that we always coded the entire conflict year from January 1 until December 31.

Angola (1994-2003)*: Africa News, Associated Press Worldstream, BBC Monitoring, Financial Times, Global News Wire,* and *The New York Times* (Marian Schmidt, Sandra Geldmacher, Robert Bauer).
Angola-Cabinda (1994-2003): *Africa News, Associated Press Worldstream, BBC Monitoring, Financial Times, Global News Wire,* and *The New York Times* (Marian Schmidt, Sandra Geldmacher Robert Bauer).
Azerbaijan (1990-1996): *Associated Press Worldstream, BBC Monitoring,* and *The New York Times (*Ayla Akdoğan).
Bosnia-Herzegovina (1990-1995): *Agence France Press, Associated Press Worldstream, BBC Monitoring, The Christian Science Monitor, CNN, Financial Times, Global News Wire, The New York Times, United Press International,* and *Washington Post* (Vanessa Dinter).
Chad (2004-2008): *Africa news, Associated Press Worldstream, BBC Monitoring, The Christian Science Monitor, Financial times, Global News Wire, The New York Times,* and *United Press Worldstream* (Roman Kropp).
Chechnya (1993-2006): *The New York Times (*Elina Brutschin, Simone Günther, Juliane Krüger, Lukas Nieslony).
Colombia (1999-2003): *Associated Press Worldstream,* and *BBC Monitoring* (Johanna Hartung, Anna Hölzer).
Côte d'Ivoire (2000-2006): *Africa news, Agence France Presse, Associated Press Worldstream, BBC Monitoring, The Christian Science Monitor, Financial Times, Global News Wire, Intelligence Online- Française, Le Figaro, The New York Times,* and *United Press International* (Tobias Schill).
Croatia (1990-1996): *Agence France Press, Associated Press Worldstream, BBC Monitoring, The Christian Science Monitor, CNN, Financial Times, Global News Wire, The New York Times, United Press International,* and *Washington Post* (Vanessa Dinter).
Democratic Republic of the Congo (1995-2002): *Africa News, BBC Monitoring,* and *The New York Times.* Coded by Marius Bayer (01.05.08 - 31.08.08) (Kathryn Claycamp, Benedikt Göller, Christiane Siegert).
Liberia (1989-1994 and 2001-2005): *Africa News, BBC Monitoring,* and *The New York Times (*Gwendolin Aschmann, Julia Koch, Simone Günther).
Macedonia (1998-2001): *Associated Press Worldstream, BBC Monitoring,* and *The New York Times* (Ann-Kathrin Hess, Alexander Bräunig).
Niger-Eastern (1994-1999): *Africa News, BBC Monitoring,* and *The New York Times* (Diana Virgilio, Johannes Erhard).
Niger- Air and Azwad (1990-1999): *Africa News, BBC Monitoring,* and *The New York Times* (Diana Virgilio, Johannes Erhard).
Republic of the Congo (1991-2005): *Africa News, BBC Monitoring,* and *The New York Times* (Tim Knickel, Naomi Bosler).
Serbia (1998-2001): *Associated Press Worldstream, BBC Monitoring,* and *The New York Times* (Vanessa Dinter, Ann-Kathrin Hess, Marjenn Witt).
Slovenia (1991): *Associated Press Worldstream, BBC Monitoring,* and *The New York Times* (Vanessa Dinter).
Sudan (1953-2007): *Africa News, BBC Monitoring,* and *The New York Times* (Vita Thormann).

<u>Sudan-Darfur</u> (2001-2008): *Africa News, BBC Monitoring,* and *The New York Times* (Vita Thormann).

<u>Uganda</u> (1979-2002): *News, BBC Monitoring,* and *The New York Times* (Robert Bauer, Lena Bringenberg, Sandra Lutz, Jana Merkelbach, Eva-Maria Niedermeier, Tobias Schill).

Appendix C: Comparison with other data sets

Comparisons with the data sets from UCDP are not that straightforward. Sometimes there is no or little overlap on the selected years (see Croatia or Azerbaijan) or regarding the actors included (see e.g. Congo-Brazzaville). Furthermore, UCDP one-sided violence dataset coded per actor and thus the figures sometimes contain casualties in several countries, and not as in KOSVED differentiated by country. We made a comparison of selected cases with the UCDP One-sided violence dataset v1.3_2011 (Eck & Hultman 2007) as well as with the UCD GED data for the African conflicts (Sundberg et al. 2010).

Comparison with the UCDP data on one-sided violence v.1.3_2011

Croatia: For the conflict for independence in Croatia, the Eck/Hultman data only coded OSV for 1993 (28 victims, best estimate) and 1995 (85) as violence taking place in the location "Croatia" and the only actor being the Government of Croatia. KOSVED, on the other hand, counts deaths for the years 1990-1995 with violence committed by Croats and by Serbs on Croatian territory. The peak of violence takes place in the KOSVED data in 1991 with 285 civilians killed by Croats and 617 by Serbs. 1991 was also the year where the conflict was at its peak and where several massacres were reported.

Azerbaijan: The UCDP one-sided violence dataset data only provides one observation, 161 civilians killed (best estimate) by the actor "Republic of Nagorno-Karabakh" in 1992. KOSVED lists casualties for the years 1990-1996, committed by the Armenian and Azerbaijani side. For example, for the year 1992 KOSVED reports 471 killings by Azerbaijani and 562 killings by Armenians. In 1991, there are 37 and 72 victims respectively reported. The total number of civilians killed as reported in KOSVED is 1409 and is closer to the reported 1264 Armenian citizens and unknown number of Azerbaijani civilian victims according to Wikipedia.

Comparison with UCD GED data for African conflicts

For the conflicts in Africa, we concentrated on a comparison with the UCDP GED data. Again, there are considerable difficulties for a direct and systematic comparison of all cases, in particular with the different actor constellations. For this reason we concentrate on the total number of killings and chose arbitrarily individual conflict actors that were easily identifiable in both data sets. First we present correlations of the logarithm of total killings by all conflict parties in the two data sets aggregated on a weekly or monthly level in the table, complemented with some additional information on each conflict in the text below.

Pearson Correlation Coefficients of all killings

	weekly	monthly
Angola	.257 (n=512)	.225(n=119)
Chad	.326 (n=213)	.688 (n=49)
Congo-Brazzaville	.099 (n=213)	.229 (n=49)
Congo-Kinshasa	.150 (n=402)	.243 (n=93)
Ivory Coast	.192 (n=113)	.783 (n=27)

Liberia I (1989-1996)	.287 (n=362)	.592 (n=85)
Liberia II (2001-2005)	-.042 (n=127)	-.129 (n=30)
Niger	.415 (n=427)	.404 (n=99)
Sudan	.217 (n=1025)	.509 (n=233)
Uganda	.435 (n=362)	.678 (n=84)

In many conflicts the two data sets show a high agreement in the weeks of none-events, whereas the agreement in the few weeks or month with high casualties is not as high.

Angola*: The correlation is not very high between the two data sets. In the following we describe the differences. As a matter of fact in 358 out of 512 weeks both datasets report no casualties. The summary statistics and standard deviation of KOSVED are higher with 6.326172 (SD=27.29869) than in UCDP data with 3.673828 (SD=16.1568). KOSVED reports more than 100 casualties in 11 weeks, whereas UCDP only does in 2 weeks. The differences are especially striking that in 46 weeks KOSVED reports on average more than 40 casualties when UCDP has none at all. In turn the reverse holds for 77 weeks where UCDP reports on average more than 15 victims when KOSVED has zero. There are only 31 weeks for which both data sets report killings but even for these weeks there is a very small correlation of the log(killings) of r=.07. When aggregated on the monthly level, there are in 56/119 month no casualties in either data. In the 13 weeks when both datasets report one-sided violence the correlation is higher with r=.42. Considering one specific conflict actor, for example UNITA, on the weekly level the correlation between the log (OSV) by UNITA is .27. For the OSV by government forces the correlation is even negative. However, there are so few events of reported killings by government troops that this is not surprising. The summary statistics for these actors in both data sets are not that far apart. If aggregated to the yearly level, there is still no rapprochement between the two datasets. In particular the reported violence for the years 1998 and 1999 are very high in KOSVED (958 and 1013) and relatively low in UCDP GED (62 and 350), whereas in 2001 UCDP GED reports more than 1000 casualties and KOSVED 208. Based on conflict developments in this case the UCDP trend seems more plausible.

* For the purpose of comparison we merged the Angola and Angola-Cabinda data together.

Chad: For Chad the correlations between the two datasets fare better. KOSVED has again a much higher mean and standard deviation. Both datasets agree for 178 (out of 213) weeks that there are no casualties reported. For the 6 weeks where both report casualties then correlation is r=.98. KOSVED reports killings in 19 weeks when UCDP reports zero victims, whereas the reverse is reported for 9 weeks. In 32 out of 49 months there is no one-sided violence reported in either data set. In only 9 month both report casualties that are correlated at r=.42.

Congo-Brazzaville: In 181 weeks both datasets report no casualties. Only in 3 weeks both report one-sided violence. UCDP reports one-sided violence for 18 week and KOSVED for 11 week when the other dataset reports zeros. Both datasets report 100 or more casualties in one week only but the weeks differ. The differences in the actors that are coded are one explanation for the discrepancies. KOSVED has many events for Ninja, a conflict actor that does not appear in the UCDP GED data. UCDP GED, on the other hand, has many events for Ntsioulous, a rebel

group that emerged after 1998, thus is out of the time range coded by KOSVED. Both data sets have few casualties and events for Cobras (UCDP several events with very few victims, in KOSVED 1 event with 20 killings & 1 event with 5 injuries). For the Cocoye rebel group, there are no events in UCDP data and very few in KOSVED about harms and few deaths.

Congo-Kinshasa: In 202 (out of 402) weeks both datasets report no casualties. The correlation in the weeks when both report casualties is with r=.16 not very high. UCDP GED reports victims in 92 weeks when KOSVED reports none, whereas KOSVED has coded victims in 61 weeks when UCDP has zeros. UCDP GED has a much higher mean than KOSVED (88.53 vs. 33.37) with more than 5000 victims in a week, whereas KOSVED reports at the most 900 (as its most conservative estimate). If we take the violence attributed to one specific actor, for example the AFDL, the correlation is not much higher (r=.23).

Ivory Coast: In 84 (out of 114) weeks both report no one-sided violence. In 10 weeks UCDP reports OSV and in 14 weeks KOSVED while the other dataset reports zeros. Only in 4 weeks are casualties reported by both datasets. Sometimes the discrepancies can be explained by a slight delay in the reporting. For example, in a couple of weeks KOSVED reports a certain number of killings that are reported by UCDP for the week later. Thus, if the data is aggregated on the monthly level, the correlation is much higher.

Liberia I: Of the 362 weeks for 274 weeks no casualties are reported in either dataset. For 21 weeks KOSVED reports one-sided violence while UCDP doesn't, for 49 weeks it is the reverse where UCDP has reports and KOSVED does not. In 17 weeks both report one-sided violence and then there is a correlation of r=.32. Both datasets reveal more weeks with higher casualties (7 weeks in UCDP and 6 weeks in KOSVED with more than 100 victims per week), however, not with a temporal overlap.

Liberia II: The correlations reported are negative for this conflict. There is an overlap of no one-sided violence reported in 103 week and the same number of victims reported in one week. Otherwise UCDP and KOSVED report casualties in 11 and respectively 12 weeks when the other has none. Sometimes this appears to be just a short time lag, other times there is no apparent explanation. KOSVED reports one week with 100 casualties in 2002, whereas UCDP reports 2 weeks with more than 100 victims in 2002.

Niger: Although both datasets provide many weekly observations, the vast majority of these are zeros. In 418 of 427 weeks both datasets report no one-sided violence. UCDP reports casualties in two weeks only, and KOSVED in 8 weeks.

Uganda: The correlations for weekly killings by the LRA is somewhat higher (r=.54). In many weeks we can observe one-sided violence. Only in a third of the weeks (130 of 362) both datasets report zero victims. In 96 weeks both report one-sided violence with a correlation of r=.39 during these weeks. In 85 weeks KOSVED reports OSV when UCDP doesn't and UCDP in 41 weeks when KOSVED reports zero one-sided violence. There are many casualties in this conflict but not very often weeks with massive numbers. All weeks in KOSVED are below 100 victims and in UCDP there are two weeks with more than 100 victims.

Sudan*: About three fourth of the weeks are reported in both data sets as zeros. Only in 52 weeks do we have violence reported by both data collections with a quite low correlation of r=.07. For

96 weeks KOSVED reports violence and for 114 weeks UCDP while the other data set records zeros. While both datasets report several weeks of excessive violence (21 weeks in UCDP and 15 weeks in KOSVED with more than 100 victims), they do not correspond in the timing. If aggregated to the monthly level the correlation is again better. The summary statistics of the two datasets are quite similar with UCDP having a mean of 6.81 (SD 34.17) and KOSVED a mean of 5.99 (SD 46.05).

*Although we present our data separately for the conflict in Darfur and in the rest of Sudan, for the purpose of this comparison we merged the total number of victims per week and month together.

Altogether, there are quite many differences in coding decisions that were made by both data sets despite the apparent similarity, in particular which conflict years to include (i.e. UCDP for now only includes years in which there are more than 25 casualties) and apparently also differences about the conflict actors that were coded with KOSVED being less limited.