



Coding one-sided violence from media reports

Cooperation and Conflict
48(4) 556–566

© The Author(s) 2013

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0010836713507668

cac.sagepub.com



Sabine Otto

Abstract

Event datasets on political violence, which are comprised of coded collected news reports, have enjoyed a renaissance within the academic community. The inclusion of civilian fatalities within these datasets is a promising and welcomed advancement regarding the availability of data on one-sided violence. However, these datasets are often criticised due to their heavy reliance on media records, which may be tainted by biases. So far, little attention has been paid to the specific problems that arise in the coding procedure with respect to one-sided violence. This article addresses such difficulties by discussing particular challenges presented by media biases and by providing empirical evidence from coding one-sided violence. Furthermore, solutions and strategies are offered to the issues that could affect the coding process, including increased transparency, definition-adaptation, and the use of appropriate statistical models.

Keywords

Civil war, event data, one-sided violence

Introduction

Only recently conflict researchers have started to investigate the occurrence and patterns of one-sided violence – the purposeful use of violence against civilians during armed conflict. Although scholars disagree over why governments and rebels intentionally target civilians during conflict, they agree that the use of one-sided violence serves as a strategy to reach certain goals (cf. Downes, 2004; Hultman, 2007; Kalyvas, 2006). Due to data limitations, the research has been carried out using different, often conflict-specific datasets and by applying divergent, perhaps incompatible, definitions of one-sided violence. Thus, due to different coding procedures the results are not easily comparable and it remains unknown whether case-specific explanations hold across conflicts. A promising and welcomed advancement is the inclusion of civilian fatalities into recently created event datasets on political violence.¹ These data may help overcome previous limitations and allow researchers to study systematically civilian victimisation

Corresponding author:

Sabine Otto, Department of Politics and Public Administration, University of Konstanz, Room Y 325, Box 216, 78457 Konstanz, Germany.

Email: sabine.otto@uni-konstanz.de

within and across conflicts. Event databases on political violence register individual conflict incidents by their spatial and temporal coordinates. Each event contains the information on who did what, when, to whom, and where. Event data are therefore a powerful tool for analysing in a detailed and disaggregated manner the violent behaviour and interactions of armed actors.

Recently created event datasets are often criticised for their heavy reliance on media records. Due to inherent biases in news reports, inferences drawn from these datasets can result in misleading outcomes (e.g. Davenport and Ball, 2002; Earl et al., 2004; Gerner and Schrod, 1994; Hawkins, 2011; Smith et al., 2001). Political scientists have been dealing with this issue for decades (cf. Doran et al., 1973), but only recently have peace and conflict researchers rekindled the topic of media bias (Chojnacki et al., 2012b; Öberg and Sollenberg, 2011; Schrod, 2012). Nevertheless, little attention has been paid to the specific problems arising from coding one-sided violence. Definitions of civilian targeting are complex, which creates empirical challenges in terms of observability and measurability. To accurately code this type of violence, a certain set of precise information is required that might not be available or that might be erroneously contained in media reports. This raises several questions. What are the concrete challenges to coding one-sided violence? What kind of definition can realistically be operationalised? What steps can be taken in coding one-sided violence from media reports in order to mitigate common biases and pitfalls? Answers to these questions are highly important, for both the providers and users of event datasets. Being aware of these issues will help improve the coding of one-sided violence, thus adding to its reliability. Knowledgeable use of these data also prevents making precipitous conclusions regarding civilian victimisation in armed conflicts, an agenda which is still emerging and highly politically relevant.

This article will address challenges to coding one-sided violence from news records, and propose strategies to mitigate against the most common biases and pitfalls. It proceeds as follows. After discussing the promises of media-based event datasets – as opposed to data collected by other means – I elaborate on the specific challenges that arise from doing so. Next, I present results from an analysis of one-sided violence, thereby demonstrating such challenges in practice. I conclude with strategies on how to cope with media bias.

Different methods to collect data on one-sided violence

Research on one-sided violence has been carried out based on a wide range of data sources and different ways of coding this type of violence. Generally, four types of data collection can be distinguished: surveys, official records and archives, data from non-governmental organisations (NGOs), and news records. An example of using survey data is Humphreys and Weinstein's (2006) work. The authors interviewed more than 1000 ex-rebels in Sierra Leone to explore the relationship between internal characteristics of the faction and level of civilian abuse. Civilian victimisation is constructed as an index of everyday policies and practices by groups reflecting the abusive or cooperative relations with the civilian population (Humphreys and Weinstein, 2006: 205). Kalyvas (2006) uses official records from the Greek civil war. Based on the detailed information

provided by court records and village civil registries, he distinguishes selective and indiscriminate violence perpetrated by governments and rebels depending on the degree of territorial control. Data collected by the Truth and Reconciliation Commission are the primary source of Fielding and Shortland's (2010) investigation on changing levels of abuse against civilians during the Peruvian civil war. In their definition, civilian abuse comprises the use of lethal as well as non-lethal violence. Wood (2010) uses the Uppsala Conflict Data Program (UCDP) data on one-sided violence, which are based on news reports as well as on NGO and United Nations reports. Applying the definition provided by UCDP he explores the connection between rebel capacity and the intensity of violence committed against civilians.

Event datasets established on media records are based on a certain selection of news sources. Nowadays, a large variety of newspaper and news wires are easily accessible via data services such as LexisNexis or Factiva, which are also possible to download. Technical developments have made it possible to support human coders with a variety of technologies to extract information in a standardised manner and to enhance the speed of coding (cf. Leetaru and Schrodt, 2013). These datasets are promising for three reasons. First, news records provide a vast pool of detailed information about conflicts from all over the world, allowing coding of one-sided violence for all conflict-affected countries. Second, event datasets based on media outlets can be created at relatively low cost and in a highly time-efficient manner. Third, due to constant news coverage, real-time coding is possible, thereby allowing datasets to be updated on a regular basis and enabling research on a very recent occurrence of civilian victimisation. Due to these benefits, event data methods are in renaissance. The critical literature on media-based event data is not suspicious of the technical aspect of the collection procedure, but rather of the quality of the media reports. Scholars have pointed out that news records suffer from inherent biases that present certain challenges to the collection of data and their subsequently analysis. Henceforth, I will discuss particular challenges that are relevant to the coding of one-sided violence.

What are the challenges of coding one-sided violence based on media data?

Although event datasets based on media outlets are a promising development, certain challenges arise when coding one-sided violence. Scholars usually distinguish between two groups of media bias (cf. Earl et al., 2004: 67). *Selection bias* refers to the selectivity of event reporting by newspapers and news wires. *Description bias* has to do with incorrectly and incompletely reported information on events. The former can be a consequence of low media attention. Scholars have pointed out that many conflicts and the most deadly among them are often not reported (cf. Hawkins, 2011). Selectivity results further from unequal news coverage in urban and rural areas (proximity bias) resulting in an under-reporting of incidents in remote areas. Furthermore, media attention depends on the victim's status and the number of casualties. On the one hand violence committed against an 'important' person, such as a politician, businessman or foreigner, is seen as more 'newsworthy' than the killing of an 'ordinary citizen' (significance bias). On the other hand incidents of large-scale civilian abuse are more likely to be reported than

events of single-civilian victimisation (intensity bias, cf. Earl et al., 2004; Moeller, 1999; Mueller, 1997).

Description bias refers to the veracity and the completeness of a reported event (McCarthy et al., 1999). Erroneously reported information can result from the misrepresentation of facts. Reporting may be prone to incorrectness when the information is obtained from parties to the conflict who have an interest in spreading certain information (source bias). Specifically, fatality counts and the type of victim are sensitive to bias because they are often the subject of conflict actors' propaganda (Chojnacki et al., 2012b). The omission of information refers to events that are described incompletely. The level of available event details varies widely across news reports (Abbott, 2006). Certain pieces of information that are necessary to enable accurate coding of the event of interest might not be included in the news report (omission bias). In order to empirically measure and to precisely code one-sided violence, a certain set of information is required from the news record. First, it must be known whether or not the perpetrator is a member of an organised armed group in order to point out the difference between conflict-related fatalities and criminal or private violence. Second, the identity of the target – whether it is a civilian – is vital.² Third, information about the context is essential to distinguish civilian targeting from those civilians who were accidentally killed in battle activities (e.g. crossfire). Articles that do not provide this minimum set of information increase the probability of coding the wrong type of violence. In the next section I exemplify how these groups of biases influence the coding of one-sided violence in practice.

Coding one-sided violence in practice

To illustrate whether and how media biases impact the coding of one-sided violence, I present the results of an analysis scrutinising the UCDP Georeferenced Event Dataset (UCDP GED) on Afghanistan for the year 2009. However, it must be noted that the data are a preliminary version, and the data may differ in subsequent revisions.³ In particular, the examples presented here may not apply to the same extent in cases where other data sources (e.g. NGO reports) were used by UCDP. I have chosen the UCDP GED data for several reasons. UCDP GED data will eventually have global coverage and the data are compatible with other datasets released by UCDP. Furthermore, UCDP GED systematically includes data on one-sided violence unlike, for example, the Armed Conflict Locations and Events Dataset (ACLED, Raleigh et al., 2010). Moreover, UCDP is one of the few data providers whose data are accessible to the public and updated regularly. Thus, UCDP GED data are one of the best documented and most prominent event datasets that include one-sided violence.

UCDP defines one-sided violence as 'the use of armed force by the government of a state or by a formally organised group against civilians which results in at least 25 deaths per year'. In order to isolate civilian fatalities as purposeful acts, UCDP provides a narrow but rather ambitious definition of one-sided violence.

The concept of one-sided violence encompasses only those fatalities that are caused by the intentional and direct use of violence. Intentional killings refer to any action that is taken to

deliberately kill civilians. ... Direct killings encompass all deaths caused directly by an actor, such as by bombing or shooting. (Eck and Hultman, 2007: 235)

What matters in distinguishing civilian casualties as a 'by-product' of military interactions from intentional killings is 'the intended target or intention of the attacking forces' (Sundberg, 2008: 3). Intentionality does not refer to a judicial assessment, nor to the knowledge of an actor's motivation. Rather, UCDP examines whether an event can be coded as one-sided violence on the basis of several criteria: the plausible civilian target, the means by which the killing was carried out, presented evidence, and credible statements of guilt. Hence, intentionality is often inferred from the violent act per se if a possible military or conflict-related target is absent (Hicks et al., 2011).

My investigation based on the UCDP GED data attests that two types of biases are especially likely to lead to distortions in event datasets created from news sources. I found valuable indications that significance bias strongly influences the reporting about certain types of victims. I also discovered that omission bias is a formidable challenge when coding one-sided violence, which results from the mismatch between definitions of one-sided violence and the information that is provided by the news records.

How can the aforementioned media biases play out when coding one-sided violence in practice? Identifying selectivity is neither an easy task, nor is it fully possible to resolve. News reports only stand for a subsample of the true population of violent events. Whether and how much the sample represents the population of all events of one-sided violence, however, remains unknown (Earl et al., 2004: 70). To uncover significance bias the information about a victim's status must be mentioned in the article, such as: 'Spokesman... told Afghan Islamic Press that the Taleban abducted a member of the council of Janikhel District and tribal elder, Shah Mohammad Zhewandi, from his house' (BBC February 9, 2009). My analysis of the types of victims, however, shows that in 64% of all events (35 out of 55) the victims were public persons working as (former) local politicians and state officials or employees of major companies.⁴ In 36% of all cases the victims were unspecified or they were 'ordinary' people killed in a private setting. This indicates that it is more likely that the killing of a representative of wider public interest makes it into the news than the killing of normal citizens (significance bias). Identifying proximity bias is more challenging, because precise locations are often not reported in media outlets – as in the following example: 'Insurgents killed a religious leader overnight in Paktia [province], the education ministry said' (Reuters June 14, 2009). In 69% of all cases the occurrence of one-sided violence is described with vague location details such as district, province or parts of the country. In all cases coded by UCDP GED, only 31% of the news articles provide a precise location as a town or village name or at least a limited area around a specific location.⁵ It is impossible to conclude whether more cases of one-sided violence were reported from urban areas as suggested in the literature. The lack of exact locations in media records, nevertheless, limits research on a greatly disaggregated level.

The most challenging difficulty when coding one-sided violence derives from omission bias. An ideal and unambiguous example of one-sided violence according to the UCDP definition is the following: 'Taliban insurgents kidnapped and beheaded two Afghans, including a tribal elder, on Sunday for "aiding the government"' (Reuters

November 16, 2009). Nevertheless, news records often do not provide (stated) intentions. Only 25% of all cases contain the actors' intentions. In the majority of cases the target and evidence of the direct use of force are used instead to classify the type of violence. The following, for example, is coded as one-sided violence: 'The Konar police chief, in other report, told [AIP] that the Taleban shot a husband and wife in Watapor District of the province this morning' (BBC September 7, 2009). While no intention is mentioned in the article, the targets are civilians and were directly killed. However, often the reported information might not provide enough evidence to code an incident as one-sided violence. For instance: 'On Sunday, a bomb on a bicycle exploded in a busy market in Gereshk district of southern Helmand province, killing four' (AP May 3, 2009). From the context it could be inferred that the bomb was intentionally placed in a social meeting point – the market – in order to kill a high number of ordinary people. But it is difficult to verify intentionality. The problem of uncertainty over intentionality has been discussed earlier (Hicks et al., 2011). When there is only weak evidence of intentional killing, UCDP's practice is to place these casualties into the 'high' estimate category instead of into the 'best' one. Such cases might be re-evaluated if further proof shows up. Users should keep this in mind and run their analysis with the different estimates to increase the robustness of their results.

A general problem with coding one-sided violence is that the perpetrator is unknown. Often, the only fact that is known is that a violent act against civilians was committed: 'According to reports, unknown gunmen killed Mowlawi Mohammad Rasul, member of the Ulema Council of Kandahar, in the Karez Bazaar area of Kandahar town last night' (BBC January 6, 2009). While the context can provide sufficient information to be sure that an event was not a private murder but rather political violence, assigning acts of violence against civilians to a specific actor is problematic when the perpetrator is unknown. In 47% of all one-sided violence cases coded by UCDP GED the perpetrators were unknown, but the Taleban was coded as the responsible party. Especially problematic is one case where a massive bomb attack in Kandahar in August 2009 killed an estimated 40 civilians. In two of three articles used by UCDP to code this event, it is explicitly stated that nobody had taken responsibility. In the third article the Taleban denied any responsibility. Notwithstanding, this event is coded as one-sided violence committed by the Taleban.⁶ However, if the actor's attribution in such cases is coded through inferences based on the coder's expertise (Eck, 2012: 135), these cases should be flagged. Coding decisions derived from the coder's knowledge may introduce a new form of bias and may lead to limited transparency.

A related issue is that news reports often only contain the information concerning killed civilians, and not the perpetrator, the target or intention, nor any detailed context information. As an example: 'A roadside blast killed a woman in western province of Herat on Sunday, a provincial official said' (Reuters March 9, 2009). Or: 'In the central Ghazni province, a roadside bomb killed four construction workers' (AP April 3, 2009). Given little information about an incident, several different scenarios are possible. The coder does not know whether these acts of violence were deliberately carried out in order to harm civilians. In the whole sample of violent events for the year 2009 in Afghanistan at least 50 events are incompletely reported, containing only the information that roadside bombs exploded and civilians died. UCDP GED

codes these events as battle-related deaths, because it cannot be ruled out that a military target was intended (Hicks et al., 2011). It should be kept in mind that this is a rather conservative coding procedure, which may result in undercounting fatalities from one-sided violence.

This brief analysis has shown that media bias is a serious challenge for coding one-sided violence. The impact of each bias unfolds differently, however. Blaming media records as the only source of distortion is too easy. Difficulties also arise from the transformation of source data into event data. If source data have the potential for uncertainty, clear documentation of how to deal with this uncertainty is indispensable to ensure transparency in the coding procedure, to determine intercoder reliability, and to control for intersubjectivity. Despite rather extensive documentation the UCDP GED codebook may benefit from some more elaboration and illustration on the coding rules for one-sided violence. The codebook's appendix comprises very detailed coding rules for event types as well as for temporal and geo-precisions, which capture a variety of uncertainties.⁷ A similar section for the different types of violence might be helpful for the user to understand easily what is measured by the data and to estimate how much uncertainty the data encompass. In the following section I offer some strategies to help creators of event datasets and researchers mitigate the aforementioned challenges.

Tentative solutions: How to mitigate against media bias

Event datasets on political violence based on media records are a promising development for investigating within and across conflicts why governments and rebels victimise civilians. They provide the opportunity to open the black box of civilian targeting by allowing the analysis of temporal and spatial variation in incidences of one-sided violence. Despite the promises, however, event data entail certain challenges. Selectivity in reporting and how events are described impact the coding of one-sided violence. While both groups of biases are not fully resolvable, certain strategies can mitigate their impacts. To moderate selection and description bias, different measures could be taken by the creators of event datasets and their users. This includes increasing transparency of the coding procedure, adopting common definitions of one-sided violence, and using the appropriate statistical models.

First, providers of event data can help cope with description bias if they design the coding process as transparently as possible. Codebooks should contain all different types of political violence and the corresponding coding rules. In addition, concrete examples of how certain types of events are coded would help scholars understand what they can expect from the data. Making coding documentation available is another way to allow researchers to retrace coding decisions. News records often describe events of one-sided violence poorly, which produces uncertainty over whether and how one should have coded an event. One alternative would be to flag those cases where there is a lack of correspondence between the definition and data source, or if the source of information seems to be unreliable. This would give the users a choice over how much uncertainty they are willing to accept in their analysis. Another suggestion is to apply a rule about how much information a news record has to contain in order for it to be used as a coding basis. If this criterion is not met, an event will not be coded.

Second, coding civilian targeting from media reports often suffers from mismatching between the coding definition and the information provided by the news record. The gap between theoretically meaningful definitions and empirically measurable ones should be kept in mind by data providers and data users. Certain types of theoretically well-founded and demanding definitions of one-sided violence might not be addressed when using media reports. One option is to emphasise this gap by revealing the extent to which the data resemble the theoretically developed definition. Another option is to avoid over-complex definitions. Other data projects, such as Event Data on Armed Conflict and Security (EDACS) and the Konstanz One-Sided Violence Event Dataset (KOSVED), apply less demanding definitions of one-sided violence compared to the one provided by UCDP. Both refrain from including intentionality in their definitions. Of course, this step decreases the precision of definitions but it improves the matching of the definition with what is actually coded. EDACS differentiates between fighting and one-sided attacks, where the latter are defined as 'direct unilateral violence by organized groups aimed at civilian or military targets' (Chojnacki et al., 2012a: 4). EDACS separates the target and the type of violence, which allows one to differentiate fatalities resulting from one-sided military attacks from battle-related deaths and civilian victimisation. The incorporation of unintentional killings of civilians, however, limits research on strategies of one-sided violence using EDACS. KOSVED, in contrast, defines one-sided violence as 'violent acts that an organized group directs against non-combatants and that result in the immediate physical harming or death of more than one individual' (Schneider and Bussmann, 2013: 638). Distinguishing between civilians as 'primary targets' (intentional killing) and as 'secondary targets' (unintentional killing) allows one to include both forms of violence separately without losing relevant cases.

Third, researchers have to make themselves familiar with the data, not only to figure out what kind of research question can be answered with the data, but also to choose appropriate statistical models. Data based on media outlets incorporate uncertainty. The true population of one-sided violence is unknown. Absence of one-sided violence can either be due to a 'true' absence of violence or to a 'false' absence because of missing data (the news does not report all events). Different statistical models permit researchers to incorporate this uncertainty (heterogeneity) into their analysis. Split-population models, for example, allow one to model unobservable phenomena and to correct for measurement bias. These models incorporate the assumption that cases, which take on a single value, are composed of two populations. Each population can be identified conceptually but not empirically (cf. Beger et al., 2011). Multiple systems estimation, in contrast, aims to come as close as possible to a true population of events by drawing on various data sources to estimate the total number of fatalities resulting from one-sided violence. Matching cases that are available in more than one data source enables one, firstly, to model the process by which violent incidents are disclosed. Estimation of the number of unrecognised cases follows in a second step (cf. Guberek et al., 2010; Lum et al., 2010).

At the end of the day, it is possible to code one-sided violence from media records. Event data provide a great opportunity to investigate why and when civilian victimisation occurs during armed conflict. If media bias is taken seriously by coders as well as by researchers, and if both parties use appropriate means to mitigate it, the distortion of

event data from media bias can be kept to a minimum. A cautious approach to creating and using event data will not only avoid cascading errors but will enhance progress in a still emerging research agenda.

Acknowledgements

The author would like to thank Nils Weidmann for extremely useful advice and comments. The author also thanks Ralph Sundberg and the UCDP team for generously providing a preliminary version of UCDP GED data on Afghanistan 2009. The author is also grateful to the Peace Research Institute Oslo, especially Håvard Strand, for hosting me from February 2012 to June 2012.

Funding

This research has been made possible due to a E.ON Ruhrgas scholarship and a GSDS scholarship from the University of Konstanz

Notes

1. For example: ACLED – Armed Conflict Locations and Events Dataset; CERAC – Datos del Conflicto Armado en Colombia; EDACS – Event Data on Armed Conflict and Security; KOSVED – The Konstanz One-Sided Violence Event Dataset; or UCDP GED – Uppsala Conflict Data Program Georeferenced Event Dataset.
2. The concept of ‘civilian’ has become a highly debated issue in violent settings where lines between government forces, rebels and population are blurred. An extensive discussion has been done elsewhere (cf. Armbrorst, 2010; Gross, 2009; Valentino et al., 2004). Codebooks of datasets usually follow a more pragmatic and less theoretical rigorous definition whose core argument is being an unarmed person at the time an event occurs.
3. The codebook provided for the preliminary data version is: *The UCDP Georeferenced Event Dataset (GED) Codebook Version 1.0*, draft as of March 2010, elaborated by Sundberg et al. (2010). Revised codebooks have been checked for major changes.
4. UCDP coded 56 events of one-sided violence but one event was coded twice.
5. Specific location or a limited area around a specific location refers to the geo-precision coding rules 1 and 2 taken from the UCDP codebook (Sundberg et al., 2010). Regarding all events of political violence only 24% of all cases provide specific locations or limited areas around specific locations.
6. For more details see the following articles. Reuters, August 25, 2009: ‘WRAPUP 6-Massive car bomb as Afghan race runs tight’; AP, August 26, 2009: ‘5 car bombs detonate simultaneously in Afghanistan, killing at least 41, flattening buildings’; Reuters, August 26, 2009: ‘Taliban denies responsibility for Afghan blast’.
7. Event type refers to single-day, summary and continuous events (Sundberg et al., 2010).

References

- Abbott M (2006) *Dangerous intervention: An analysis of humanitarian fatalities in assistance contexts*. PhD Dissertation, Ohio State University, Columbus, OH.
- Armbrorst A (2010) Modelling terrorism and political violence. *International Relations* 24(4): 414–432.
- Beger A, DeMeritt JH, Hwang W, et al. (2011) The split population logit (SPopLogit): Modeling measurement bias in binary data. Working paper, <http://polmeth.wustl.edu/media/Paper/BegDemHwaMoo28Feb11.pdf> (last accessed 10 October 2013).

- Chojnacki S, Ickler C, Spies M, et al. (2012a) EDACS Codebook Version 3.7, The Event Data on Armed Conflict and Security (EDACS) Project, Free University Berlin.
- Chojnacki S, Ickler C, Spies M, et al. (2012b) Event data on armed conflict and security: New perspectives, old challenges, and some solutions. *International Interactions* 38(4): 382–401.
- Davenport C and Ball P (2002) Exploring the implications of source selection in the case of Guatemalan state terror, 1977–1995. *Journal of Conflict Resolution* 46(3): 427–450.
- Doran CF, Pendley RE and Antunes GE (1973) A test of cross-national event reliability: Global versus regional data sources. *International Studies Quarterly* 17(2): 175–203.
- Downes AB (2004) *Targeting civilians in war*. PhD Thesis, University of Chicago.
- Earl J, Martin A, McCarthy JD, et al. (2004) The use of newspaper data in the study of collective action. *Annual Review of Sociology* 30(1): 65–80.
- Eck K (2012) In data we trust? A comparison of UCDP's GED and ACLED conflict events datasets. *Cooperation and Conflict* 47(1): 124–141.
- Eck K and Hultman L (2007) One-sided violence against civilians in war: Insights from new fatality data. *Journal of Peace Research* 44(2): 233–246.
- Fielding D and Shortland A (2010) What explains changes in the level of abuse against civilians during the Peruvian Civil War? Working Papers 1003, University of Otago, Department of Economics, http://www.business.otago.ac.nz/econ/research/discussionpapers/DP_1003.pdf (last accessed 10 October 2013).
- Gerner DJ and Schrodt PA (1994) Machine coding of event data using regional and international sources. *International Studies Quarterly* 38(1): 91–119.
- Gross ML (2009) Asymmetric war, symmetrical intentions: Killing civilians in modern armed conflict. *Global Crime* 10(4): 320–336.
- Guberek T, Guzmán D, Price M, et al. (2010) *To count the uncounted: An estimation of lethal violence in Casanare*. Report by the Benetech Human Rights Program, 10 February 2010.
- Hawkins V (2011) Media selectivity and the other side of the CNN effect: The consequences of not paying attention to conflict. *Media, War & Conflict* 4(1): 55–68.
- Hicks M, Lee UR, Sundberg R, et al. (2011) Civilian targeting index clarification. In: *The Duck of Minerva*. Available at: <http://duckofminerva.blogspot.se/2011/10/crunching-corpse-counts-rejoinder-by.html> (last accessed 10 October 2013).
- Hultman L (2007) Battle losses and rebel violence: Raising the cost of fighting. *Terrorism and Political Violence* 19(2): 205–222.
- Humphreys M and Weinstein J (2006) Handling and manhandling civilians in civil war. *American Political Science Review* 100(3): 429–447.
- Kalyvas SN (2006) *The Logic of Violence in Civil War*. Cambridge: Cambridge University Press.
- Leetaru K and Schrodt PA (2013) GDEL: Global Data on Events, Location and Tone, 1979–2012, Paper presented at the International Studies Association meetings, San Francisco, April 2013.
- Lum K, Price M, Guberek T, et al. (2010) Measuring elusive populations with Bayesian model averaging for multiple systems estimation: A case study on lethal violations in Casanare, 1998–2007. *Statistics, Politics, and Policy* 1(1): Article 2.
- McCarthy JD, McPhail C, Smith J, et al. (1999) Electronic and print media representations of Washington D.C. demonstrations, 1982 and 1991: A demography of description bias. In: Rucht D, Koopmans R and Neidhardt F (eds) *Acts of Dissent*. New York: Rowman & Littlefield, pp.113–130.
- Moeller SD (1999) *Compassion Fatigue: How Media Sell Disease, Famine, War and Death*. New York and London: Routledge.
- Mueller C (1997) International press coverage of East German protest events, 1989. *American Sociological Review* 62(5): 820–832.

- Öberg M and Sollenberg M (2011) Gathering conflict information using news resources. In: Höglund K and Öberg M (eds) *Understanding Peace Research: Methods and Challenges*. London: Routledge, pp.47–73.
- Raleigh C, Linke A and Hegre H et al. (2009) Introducing ACLED: An armed conflict location and Event dataset. *Journal of Peace Research* 47(5): 651–660.
- Schneider G and Bussmann M (2013) *Accounting for the dynamics of one-sided violence: Introducing KOSVED*. *Journal of Peace Research* 50(5): 635–644.
- Schrodt PA (2012) Precedents, progress and prospects in political event data. *International Interactions* 38(4): 546–569.
- Smith J, McCarthy JD, McPhail C, et al. (2001) From protest to agenda building: Description bias in media coverage of protest events in Washington, D.C. *Social Forces* 79(4): 1397–1423.
- Sundberg R (2008) Revisiting one-sided violence – a global and regional analysis. Paper No.3, Uppsala University, Department of Peace and Conflict Research.
- Sundberg R, Lindgren M and Padskocimaite A (2010) *UCDP geo-referenced event dataset (GED) Codebook Version 1.0*. 5 March 2010. Department of Peace and Conflict Research, Uppsala University.
- Valentino B, Huth P and Balch-Lindsay D (2004) ‘Draining the sea’: Mass killing and guerrilla warfare. *International Organization* 58(2): 375–407.
- Wood RM (2010) Rebel capability and strategic violence against civilians. *Journal of Peace Research* 47(5): 601–614.

Author biography

Sabine Otto is a PhD student at the Graduate School of Decision Science and is affiliated with the ‘Communication, Networks and Contention’ Research Group at the Department of Politics and Public Administration, University of Konstanz. Previously, she was a visiting scholar at the Centre for the Study of Civil War, Peace Research Institute Oslo. She received my MA in Political Science from the Free University, Berlin in October 2011.