

Conflict and Gender: Data Overview

Amelia Hoover

Yale University

Benetech Human Rights Data Analysis Group

November 2009

What Do We Want to Know?

What is the differential impact of war by sex? How would we measure that? Do some types of direct violence (e.g., sexual violence) disproportionately affect women? Are women more likely than men to suffer indirect effects of war (such as hunger, disease and displacement)? How do the gendered effects of war vary from conflict to conflict?

Answers for most of these questions are surprisingly uncertain. In most wars, according to most definitions of sexual violence, women seem to suffer significantly more sexual violence than men (e.g., Conibere et al. 2005, Cibelli, Hoover and Krüger 2009). However, Leiby (2009) finds that among victims of sexual torture in Perú, violations suffered by males are often classified as torture, while the majority of females who suffer sexual torture are classified as having suffered sexual violence. This finding implies that men suffer more sexual violence than has been recognized. Given that most reported victims of torture are male, it may even imply that some forms of sexual violence are committed against men as often as against women.

Some authors have suggested that, while men are more likely to suffer direct violence during conflict, women suffer more of the indirect effects of war, including increased mortality from hunger and disease. Others have noted that the effects of war, particularly killing, fall very heavily on men, especially men of fighting age (e.g., Carpenter 2006). In nearly every conflict, men of fighting age are reported killed at a much higher rate than women of any age.

Of course, this finding depends on the assumption that reported rates of killing by sex hew closely to true rates of killing by sex. Whether this is true remains unclear. Certainly data accuracy and completeness vary dramatically from conflict to conflict. For example, Ball et al. (2007) estimate that 95-100% of killings during the civil war in Bosnia were reported, by name, to at least one source. By contrast, media and NGO lists of killings during many conflicts (e.g., Guatemala, El Salvador, Sierra Leone, Liberia, East Timor, Perú) are significantly less complete. It is unclear whether under-reporting varies by sex, however. Below, I demonstrate that female homicide victims were reported at only slightly lower rates than male homicide victims during one conflict (Kosovo, 1999).

This data overview focuses on the benefits and drawbacks of several common types of data, which must be carefully considered when these data are used to inform policy decisions. I describe in detail the work of the Benetech Human Rights Data Analysis Group toward correcting incomplete and biased data.

Types and Sources of Data

Convenience Data Data on conflict-related violence and mortality fall into four categories. Observational or “convenience” data are probably most common. Convenience data represent an attempt, however small, at enumeration of all acts of interest (e.g., all homicides; all episodes of sexual violence). Convenience data are not collected according to a survey sampling scheme. Sources of convenience data include newspaper reports, case files of human rights organizations, hospital records, truth commission testimony, and border crossing records (e.g., Ball, 2000, 2002). Convenience data have also been termed “found data.” Because these data often contain information about the victim such as gender, age or ethnicity, they often prove extremely useful in disaggregating a human rights crisis. Using convenience data, the analyst can ask: To which group (if any) was this violence targeted? Who were the key perpetrators? Which forms of violence were used most frequently?

Convenience data often are not formatted as datasets; many convenience datasets come in the form of abandoned notes or records (see, e.g., Ball 2000). However, even qualitative accounts can be coded and analyzed quantitatively. Although not all convenience data are ultimately verified, a key characteristic of a convenience dataset is that each observation corresponds to a specific victim or act. This is what distinguishes convenience data from “naïve estimates” (see below).

Even if each and every data point were verified, raw convenience data would still not necessarily indicate true patterns and magnitudes of violence. Lists produced by governmental, non-governmental, medical, journalistic or community organization are lists produced by entities with agendas, budget and staff limitations, and little or no access to some groups and areas. Social networks between and among victims also play a major role in under-reporting. Theidon (2007) documents how Peruvian communities often decided on a collective narrative to share before the Truth and Reconciliation Commission (Comisión de Verdad y Reconciliación, CVR). These decisions strongly affected the quality of the data gathered by the CVR from village to village. More generally, in any conflict situation some social groups will report to police, church, or NGO sources while others will avoid these same sources (Ball et al. 2007b).

Questions of gender, and violence targeted by gender, are especially prone to these types of data biases. For example, while rape and other sexual violence occurred frequently during the civil war in El Salvador (1980-1992) (e.g., Anaya 2007), the total number of episodes of sexual violence reported to the United Nations-sponsored Truth Commission for El Salvador (1993) was just over 300 (see below). These 300 cases represent a small fraction of the true number; many more were reported to two other data-gathering projects; still more went unreported. Deponents to early Truth Commissions often did not testify about witnessing or experiencing sexual violence, either because of the social stigma attached to sexual violence, or because the commission simply did not solicit information on sexual violence. In addition, women who report sexual violence risk ostracism and other forms of retaliation (Swiss and Jennings 2006). The accuracy of self-reported data about violence, especially including sexual violence, can

never be assumed. In many cases, uncorrected convenience data have grossly misrepresented both pattern and magnitude of violence (see Ball et al. 2007b, Lynch and Hoover 2008).

Survey Data The second major type of violence data, survey data, do not attempt to enumerate all episodes of violence. Instead, by requesting information from representative households, survey investigations extrapolate from the sample to the general population. Survey methods are significantly more expensive than convenience data, but they have a significant advantage: if the sample frame can be determined, and the sample is correctly drawn (i.e., if the sample is truly representative)

A key advantage of survey methods is their strong, well-established criteria for assessment. Given some knowledge of the sample frame (the population of interest), the sampling method and the survey instrument, debate about the validity of a survey's findings proceeds along well-established, if often technical, lines. However, knowledge of the sample frame is much more difficult when the population of interest is, for example, "noncombatant victims of human rights violations." Surveys during and immediately after conflict face a number of serious difficulties, including massive population displacement and unknown population size. Often, surveys of conflict-affected populations take place in refugee or internally displaced persons' encampments. Outside the special case in which an entire displaced population is displaced to camp settings, findings from camp surveys are often misleading when applied to non-camp displaced populations (e.g., Amowitz et al. 2002).

Indices A third type of data, known as an index, is typically derived from convenience data by combination or categorization. For example, Amnesty International and the United States Department of State each produce a five-point scale (for each country in each year) that rates countries' human rights records, relying on public accounts of the frequency of abuse. Other indices familiar in the human rights world include Freedom House and Polity scores, which attempt to combine many data points into a single scale or group of scales. In this volume, Cohen uses various sources to compile an index of sexual violence for countries in conflict [Citation to Cohen briefing paper].

While indices may be useful in distinguishing large magnitudes of violence from small, it is important to note that indices are typically derived, more or less directly, from convenience data, and that convenience data inevitably suffer from serious biases and flaws. Categorizing or otherwise transforming flawed convenience data may mask or ameliorate their flaws. Unfortunately, it also risks magnifying the inaccuracies of convenience data (see, e.g., debate regarding the Dirty War Index: Hicks and Spagat 2008; Taback 2008; Hoover et al. 2009). While many indices include a gender component, this gender information, like the index itself, relies entirely upon convenience data (and in some cases, surveys).

Naïve Approximation Naïve approximations (see Hoover, forthcoming 2010) are the fourth common source of information about the scale of human rights violations. These approximations are often based on qualitative impressions such as "many" or "hundreds" and

then extrapolated, often by simple multiplication, over time and/or space. Other naïve approximations include those derived by extrapolating (again, often by simple multiplication) survey results or other statistics to populations to which they do not necessarily apply. A naïve approximation is not a statistical estimate: there is no way to attach a statistical confidence interval to the approximation, and there exists no specific information on victims that might allow for comparison to other data sources.

An archetypal naïve approximation is the number “400,000,” frequently cited as the death toll in Darfur. In this case, the Coalition for International Justice interviewed approximately 1,100 persons in Chadian refugee camps, then extrapolated the result of this relatively small self-selected study to the full population of Darfuri refugees in Chad. In this way, the CIJ “estimated” that about 140,000 persons had died violently. The analysts similarly extrapolated a statistic from a World Health Organization survey (also collected among camp populations) to estimate that about 250,000 persons had died of indirect causes such as hunger and disease (Hagan et al. 2005, 2006).

While naïve approximations may be useful for emphasizing the magnitude of a crisis, they seldom provide reliable answers to questions of magnitude. Perhaps more importantly where questions of gender are concerned, these large round numbers provide no opportunity to disaggregate victims, violations, or perpetrators.

About the Human Rights Data Analysis Group

Surveys of conflict-affected populations are difficult and expensive. Raw convenience data represent a biased sample and often do not reflect the true situation. Yet policy-makers rightly demand data, in some form, in order to craft evidence-based systems for conflict prevention and remediation. Human rights organizations demand data in order to emphasize the scale of a problem, and in order to grasp the attention of both policy-makers and the public.

Yet neither policy-making organizations or human rights advocacy organizations can typically afford to employ experts in survey design and administration, data analysis, or statistical demography. The Human Rights Program of the Benetech Initiative provides technical assistance and analysis to human rights organizations. Its most basic project is Martus (see <http://www.martus.org>), a free software that stores human rights data on remote servers, lowering the risk that human rights organizations will lose or be plundered of important data.

Martus’ architecture is based on Ball’s (1996) model for convenience data management, known as “Who Did What to Whom?” Ball recommends that databases on conflict-related violence be designed with individual acts of violence as the fundamental unit of analysis. Every act of violence may have one or more perpetrators and one or more victims; every victim may have suffered one or more acts of violence. Data-gathering focused on “incidents” or victims of violence may suffer information loss, for example by failing to record identifying information for multiple victims of a single incident, or by failing to document multiple crimes suffered by the

same victim. “Who Did What to Whom” retains as much detail as possible while still allowing later analyses by victim or by incident.

The Benetech Human Rights Program also includes the Human Rights Data Analysis Group (HRDAG), which partners with other human rights organizations to design and implement data collection, coding, management and analysis. HRDAG projects range from simple advice regarding existing small data collections to large, ground-up database design and analysis projects. A 2007 analysis of disappearances in Casanare, Colombia utilized existing data on roughly 1,500 cases, demonstrating that in some areas, and some time periods, approximately 40% of disappearances went unreported. By contrast, its 2009 statistical report to the Liberian Truth and Reconciliation Commission documents over 160,000 acts of violence against over 18,000 victims (Cibelli, Hoover and Krüger 2009). HRDAG assisted the Liberian Truth and Reconciliation Commission in designing its statement form, training enumerators, coding the resulting data, organizing its database and analyzing the data, a three-year process.

While many HRDAG projects include data on specific violation types and victim sex, no HRDAG project has focused exclusively on the gendered impact of conflict violence. However, analyses performed for recent HRDAG partners such as the Sierra Leone and Liberia Truth and Reconciliation Commissions have included lengthy sections disaggregating reported violations by victim sex and violation type. This represents an advance over earlier Truth Commissions, which generally attended primarily to lethal violations, for which men are typically at higher risk. The Sierra Leone and Liberia commissions (whose reports were released in 2003 and 2009, respectively) continue a trend among human rights data gathering missions of increased attention to non-lethal types of violence and the indirect effects of conflict and violence.

HRDAG Projects and Methods

Data from ongoing and completed HRDAG projects represent a significant resource for policy-makers, human rights advocates and academic researchers. Several HRDAG projects utilize decades of data, collected and contributed by many partner organizations, that represent thousands, or even hundreds of thousands, of episodes of conflict-related violence. In this section I describe four such projects in detail. The cases chosen, presented in rough chronological order, reflect very different conflicts, and very different data collection strategies. I also review HRDAG’s key methodological contribution, multiple systems estimation (MSE). Finally, HRDAG’s collection of data is summarized in tabular form.

El Salvador El Salvador’s UN-sponsored Truth Commission (United Nations 1993) convened at the close of a twelve-year civil war which had been preceded by several years of intense violence against Salvadoran citizens. The Commission collected information on a limited set of violations, including killing, disappearance, torture, non-lethal physical assault, and rape, but focused primarily on lethal violations (killing and disappearance), which account for about 6,500 of the approximately 7,000 violations reported directly to the Truth Commission (nearly 95%). Of violations reported directly to the Truth Commission whose victims are named, only 30 are

episodes of sexual violence, a vanishingly low proportion in a war whose history includes frequent accounts of sexual violence (e.g., Wood 2003, Leiva 2006, Anaya 2007).

Unlike most recent truth commissions, the Salvadoran commission did not collect information on victim demographics such as age or sex. While this information can often be imputed from victim name in El Salvador, neither the Truth Commission's data nor data from the non-governmental Salvadoran Commission for Human Rights (Comisión de Derechos Humanos de El Salvador, CDHES) include explicit reference to the victim's sex. El Rescate, a Los Angeles-based human rights organization, collected and coded statements to the Legal Aid office of the Archbishopric of San Salvador. This dataset includes both explicit coding for victim sex and a wide range of violations, including both lethal and non-lethal violations.

El Rescate data provide interesting clues to what is missing from the Truth Commission data. In particular, killing accounts for only 28% of the approximately 22,000 violations coded by El Rescate. The most common violation, detention (30% of violations), is not coded in the Truth Commission data. Violations with female victims make up approximately 18% of violations reported in the El Rescate data. Most violation categories (including killing and detention, the two most common categories) include a similar proportion of female to male victims. Two exceptions are sexual assault, which is targeted primarily to women, and aerial bombardment, which is targeted to population centers rather than specific persons. Of aerial bombardment casualties with known sex, nearly half are female. Like El Rescate, data from CDHES include multiple non-lethal violation types, the most common of which is detention (36% of reported violations). More notably, killings (coded as "arbitrary execution" or "massacre") account for only about 6% of violations reported in CDHES data.

The Truth Commission for El Salvador was among the first commissions (others included South Africa and Haiti) to include an extensive statistical analysis in its final report (see Annex 2 to the Report of the Truth Commission for El Salvador). The format of the data, like the absence of data on victim demographics, reflects its age. Digital versions became obsolete following the publication of the Report; consequently, the only data remaining as of 2007 were approximately 60,000 lines of printed data residing in paper copies of the Report's Annex 2. Hoover (2008) re-digitized these data using optical character recognition software¹, but their quality remains somewhat poor. In particular, Annex 2 divided data into three main sections: (1) named victims reported directly to the Commission, (2) named victims reported to the Commission by other organizations, and (3) victims whose names were withheld.

This three-part organization poses both practical and theoretical problems. Both El Rescate and CDHES reported some, but not all, of their data to the Truth Commission. The Truth Commission's method for eliminating duplicate records is unclear, and especially unclear when a victim was reported directly to the Truth Commission *and* to another organization. Matching ("overlapping") records in two or more datasets are vital to achieving reliable estimates from

¹ ABBYY FineReader and OmniScan SE

convenience data (see below). Without knowledge of the process used to de-duplicate Truth Commission data, either across or within sections, it is difficult to model the reporting process well enough to achieve a reasonable estimate. However, because the estimation process can be automated, one way to solve this problem is by building models repeatedly, based on various sets of assumptions about the Truth Commission's de-duplication process.

In general, raw data from the Salvadoran civil war provide a confusing picture of the dynamics of violence during that conflict. They fail to answer key questions about targeting by age or by sex; the multiple definitions and categories of violence on which the three key datasets collected information further clouds this picture. It seems clear, at least, that Truth Commission data under-report non-lethal episodes of violence. If the distribution of violence type by victim sex were more clear, analysts might infer the gendered impact of this under-reporting.

Unfortunately, inferences by sex require re-coding of sex data based on victim's name for two of the three available datasets. But the methodological point here is probably more important than the outstanding empirical questions. In general, convenience data provide a partial and biased picture of patterns and magnitudes of violence. There exists better and worse convenience data (see the discussion of the Liberian Truth and Reconciliation Commission below for an example of better convenience data), but unless it is statistically corrected, the accuracy and representativeness of convenience data can never be assumed.

Multiple Systems Estimation Convenience data are cheap to produce and relatively plentiful, but they can be misleading. Beginning with assistance to the Guatemalan Commission for Historical Clarification (CEH by its Spanish acronym) in 1996, Ball and colleagues (at the American Association for the Advancement of Science and, beginning in 2003, HRDAG) applied increasingly complex demographic statistical techniques to the correction of convenience data. In the Guatemalan case, media accounts of violence failed to document an enormous spike in violence during the early 1980's, largely because the victims were indigenous Mayans without access to urban networks (such as the news media and, to a large extent, NGO's). This is a key example of the inferential errors that can occur when analysts rely on convenience data, and a key example of the importance of multiple systems estimation (MSE) to the human rights community.

Using MSE, Ball and other statistical advisors to the CEH found that approximately 200,000 Guatemalans, a large proportion of them indigenous people, had been killed during the war, information that played an important role in the CEH's finding that acts of genocide had been committed by the Guatemalan army (CEH 1996).

The intuition behind MSE is simple, although the statistics involved can be complex (see Ball et al. 2007 for a more technical explanation of the process). Consider the experiment in which two students each choose a handful of numbered balls randomly from a collection of unknown size. Each student records the numbers she picked, then replaces her selection. If the total number of balls in the jar is large, we expect the overlap between the two students' handfuls to be small

(because any given ball is less likely to appear in any given handful when there are more balls to choose from). If the total number of balls in the jar is small, we expect the overlap between the two handfuls to be larger.

This is a direct analogy to the case of human rights violations: if the number of violations is large, then organizations are less likely to document any given homicide, and the overlap between lists is likely to be small. If the overlap between lists of human rights violations can be determined, then the techniques of statistical demography can accurately estimate the number of violations that were not documented by any organization.

Statistical models of this intuition have been applied in population biology since the late nineteenth century. Researchers in demography have applied multiple systems estimation (often known as “multiple-recapture” estimation) since the publication of Chandra Sekar and Deming's landmark study of Indian natality in 1949. In both experimental studies (e.g., Burnham and Overton 1978, Chao 1992) and computer simulations (e.g., Fienberg et al. 1999), multiple recapture estimation produces highly accurate results, even for very unevenly distributed populations.

HRDAG researchers (e.g., Ball et al. 2002, 2007; Guzmán et al. 2007) have used the term “multiple systems estimation” to distinguish the use of multiple-recapture techniques in human rights applications (as opposed to, for example, wildlife management applications). The most technically demanding aspect of this work is determining the overlap -- the number of violations reported in more than one list. When the population is human rights violations, determining overlap requires very specific identifying information regarding every violation, usually including at least the victim's name, the year of the violation and the rough location of the violation. However, once the overlap is known, statistical programs can estimate the number of unreported violations--and hence, the total number of violations--very quickly (see, e.g., Baillargeon and Rivest 2007).

In a recent multiple systems project, Guzmán et al. (2007) found that the overlap between a Colombian police list of homicides and an NGO list of homicides decreased following the government's paramilitary demobilization policy (implemented in 2003). This finding implies that at least one of the two lists reported a smaller proportion of the total homicides after demobilization, and illustrated again the pitfalls of working only with raw convenience data. Lynch and Hoover (2008) conclude that MSE is the only analytical strategy that can fully account for the unevenness of conflict-related violence reporting. Although surveys generally provide better accountings than convenience data, they often fail to adequately document events that are rare on a population level, tightly clustered, or stigmatized (see Spiegel and Salama 2000, Lynch and Hoover 2008, Swiss and Jennings 1998).

Kosovo Ball (2000) analyzed several data sources on killings and migration during the civil conflict in Kosovo. Ball's 2000 report, published by the American Association for the Advancement of Science, used four incomplete data sources to estimate the true pattern of deaths

during March-May 1999: interviews with Kosovar refugees conducted by Human Rights Watch, interviews conducted by the Organization for Security and Cooperation in Europe, interviews by the American Bar Association's Central and East European Law Initiative, and a list of persons whose bodies were identified after exhumations of mass graves sponsored by the International Criminal Tribunal for the Former Yugoslavia. These data sources identified 4,400 unique deaths during the March-May period.

Based on these data, Ball et al. (2000, 2002) estimated that approximately 10,000 Kosovar Albanians had been killed during the March-May period. The 2000 report was followed by an extensive study submitted to the International Criminal Tribunal for the Former Yugoslavia (Ball et al. 2002a, 2002b); Ball served as an expert witness for the Prosecution in the trials of Slobodan Milosevic (ICTY IT-02-54) and Milan Milutinovic et al. (ICTY IT-02-57). In addition to analyzing deaths during the March-May period, Ball used a number of sources, including hundreds of pages of hand-written border-crossing records, to estimate patterns of migration during the March-May period. Ball found that the estimated patterns of deaths and migration were consistent with the hypothesis that Serbian forces had orchestrated both deaths and disappearances. Milosevic's trial ended prematurely when the defendant died; the Chamber in Milutinovic et al. found the statistical evidence unconvincing, although the defendants were found guilty of a majority of the charges against them (Hoover, forthcoming).

HRDAG data from Kosovo are notable for their narrowness: only deaths and migrations are recorded. As in the other Balkan conflicts of the 1990's, significant amounts of sexual violence were reported in Kosovo (e.g., United Nations Population Fund 1999), but HRDAG data focus entirely on deaths and migrations. The migration data lack adequate identification of victim sex, but data on deaths have surprisingly complete information on victim sex. Consequently, for the Kosovo conflict it is possible to reach a definitive and reliable answer to questions about under-reporting by sex. In particular, it is possible to determine whether the proportions of men and women in the data accurately represents the proportions of men and women in the (estimated) true number of deaths during March-May, 1999.

A key element of estimation via MSE is stratification. In order to control for unevenness, separate estimates are made for separate sectors (strata) of the population (e.g., one estimate for every geographical division, one estimate for every year, or one estimate for victims of each perpetrating group). MSE estimates have not generally been stratified by sex, largely because other strata (typically time, geography, or perpetrating group) have more frequently been discussed as causes of uneven under-reporting (e.g., Lynch and Hoover 2008). However, while other strata may be more important for achieving accurate global estimates, stratifying data by sex allows us to determine with confidence whether under-reporting varies by sex.

Kosovo: New Analysis As an illustration, I used already-matched data from the conflict in Kosovo (Ball 2000, 2002) to re-estimate the global total of deaths, stratifying by sex as well as by geography and date. Raw data from Kosovo included 4,400 deaths, of which only about 750 (17%) had female victims. Within the raw data, three of four data sources documented females at

a lower rate than males, but in no case was the difference in documentation significant.² It is particularly notable, however, that the one dataset in which the proportion of females documented was larger than the proportion of males documented was exhumation data--data that do not rely on reports by survivors. This finding could easily occur by chance. However, if it were replicated repeatedly it might indicate that women's deaths are consistently under-reported in interview data, relative to data that do not rely on witnesses, such as the exhumation data in this case.

[Kosovo table about here]

Based on the raw data, then, women do not appear to be drastically under-reported relative to men. Men are reported at slightly higher rates, but these rates could be higher by chance. However, MSE complicates this optimistic picture somewhat. When multiple systems estimates are re-calculated to account for sex, the 750 deaths of females documented in the raw data yield nearly 1850 estimated deaths, while male deaths documented in the raw data yield approximately 8300 estimated deaths. The global estimate calculated by this method is extremely close to the original global estimate reported in Ball et al. (2002).

A more important aspect of these findings is that documented female deaths account for about 40% of estimated female deaths, while documented male deaths account for about 45% of estimated male deaths. This is not a statistically significant difference, but it is a larger difference than that documented in the raw data. Differences in capture rates as measured by raw convenience data may understate the true differences in capture rates. While differences in the undercount of men and women were not statistically significant in this case, they may be significant in other cases--and this information will not necessarily be captured by raw data, even when multiple data sources exist.

Timor-Leste In Timor-Leste, HRDAG researchers helped the Commission on Reception, Truth and Reconciliation of Timor-Leste (CAVR by its Portuguese acronym) develop a comprehensive strategy for determining levels and, in some cases, causes of mortality during the Timorese independence struggle (1974-1999). In particular, because famine was a major source of mortality during the conflict, the HRDAG researchers hoped to determine the relative magnitudes of indirect and direct mortality. Silva and Ball (2006) estimated that approximately 102,800 (+/- 11,000) persons were killed during the conflict, of which nearly 20,000 were killed directly and approximately 80,000 more died of indirect causes. "Indirect mortality" is defined as deaths due to hunger and illness exceeding the number that would have been expected if hunger and illness continued at pre-conflict levels for the duration of the CAVR's mandate period.

² Essentially, the Student's T test asks whether differences in the proportion could have occurred at random. Typically, if the difference between two test proportions only happens 5% of the time by chance (according to the rules of probability), then the difference is statistically significant. None of the four differences in this case (between male and female capture proportions in raw data from the ABA, HRW, the OSCE, and exhumations) meets this criterion and hence none is judged significant.

At the time of the CAVR's inception, no lists of the dead and few vital statistics existed for Timor-Leste. For that reason, the CAVR undertook three separate data projects, including two convenience data sources and a household survey. The household survey gathered information, including victims' names and other data, from approximately 1,500 households; 7,668 statements to the CAVR; and a nationwide survey of approximately 319,000 graves, including graves dating back to the seventeenth century. After setting aside graves marked before and after the CAVR's mandate period and matching the three datasets, approximately 97,000 unique deaths were recorded between the three sets. Of these, nearly 7,500 were reported in two or more of the datasets.

The key point of interest in the CAVR data lies in the gender differences between direct and indirect mortality. Of direct and indirect lethal violations for which the victim's sex is known (approximately 87%), lethal violations to women account for 37% of the reported deaths. However, when we look at lethal indirect violations (people who died of hunger or disease due to the conflict), nearly half (approximately 44%) have female victims. Among civilians who died of direct causes, only about 11% of reported killings have female victims. This is perhaps an unsurprising finding; the hypothesis that women are more affected by indirect than direct mortality is supported by evidence from several conflicts.

A more important aspect of this finding is the fact that the proportion of indirect (versus direct) deaths, and consequently the proportion of female deaths, varies significantly between the three Timorese data sources. While Silva and Ball (2006) estimated that over 80% of about 102,000 total deaths during the conflict were indirect deaths due to famine or disease, fewer than half of the approximately 11,000 deaths reported to the CAVR were indirect deaths. The methodological implication is clear: in the case that convenience data may drastically underreport indirect deaths as a result of conflict. If (as in this case) women represent few direct casualties but a more significant proportion of indirect casualties, a higher proportion of women than men will remain unreported unless mortality surveys are used to complement self-reports.

Liberia HRDAG data from the Liberian civil war cover the entirety of the Truth and Reconciliation Commission's mandate period, 1990-2003. The data were collected to conform with a "Who Did What to Whom" data management strategy; that is, the key unit of analysis is the individual violation. Information is collected from deponents' testimonies. Each testimony may include reports of one or many violations. Each violation may have one or many victims and one or many perpetrators; each victim may have suffered one or many violations. Moreover, individual victims and violations may be reported by multiple deponents (for example, both a friend and a family member). For each violation, information about the violation date, location, type and perpetrator was collected. For each victim, information about the victim's name, age, sex, tribe, social sector, hometown and other identifying information was collected, to the extent possible.

The result is a massive dataset (nearly 180,000 violations reported), with relatively complete data about victims, violations and perpetrating groups. Only about 30% of reported violations

included information on victims' ages, and this percentage varies with type of violation. For example, killings make up about 20% of total reported violations, but only 5% of violations with known age. Victim sex data, by contrast, is available for nearly all violations (approximately 99%). About 62% of violations had a male victim; about 37% had a female victim. Of course, that pattern shifts dramatically depending on the type of violation. For example, most reported killings had male victims; most reported episodes of sexual violence had female victims.

Reported property crimes such as looting and robbery appear to reflect ownership patterns in Liberian society. Risk of property crimes appears highest among older men (relative to the proportion of older men in the population). Crimes directly related to combat (killing, torture, and forced recruitment) skew heavily male, but significantly younger. Sexual crimes generally skew toward adolescent females (as opposed to males of any age or women of socially taboo ages such as the elderly and very young). There is one major exception: the category "sexual abuse," which excludes rape but includes sexual humiliation and sexual touching, is surprisingly evenly distributed between the sexes. However, sexual abuse represents a relatively small proportion of reported sexual violations.

The graphics below represent HRDAG analysts' efforts to consider the demographic structure of Liberian society in their considerations of targeting. Because Liberia is an extremely young society (figure 1), we would expect to observe many violations against young people, even if victims were chosen randomly (rather than by targeting). When we look at the magnitude of violations by age and sex, this phenomenon can be misleading. Thus, HRDAG analyses typically include a calculation of "relative risk," the proportion of violations that occurred in a given age-sex category divided by the proportion of Liberians who fall in that age-sex category. Thus if 20% of Liberians were males aged 15-19, but 50% of all killings had male victims aged 15-19, the relative risk of killing for a male aged 15-19 is $50/20 = 2.5$. If all reported killings were targeted randomly, then we would expect all relative risk values to be 1. Figures 2a-2c show relative risk graphs for violation types that skew female. Figures 3a-3i show relative risk graphs for violation types that skew male. Figures 4a and 4b show relative risk graphs for forced displacement and all violations, respectively.

[Figure 1 about here.]

[Figures (2) about here.]

[Figures (3) about here.]

[Figures (4) about here.]

What can we infer from these graphics? Unfortunately, as with all convenience data, these represent only about the violations *reported* to the Truth Commission. In order to extrapolate from this data to all violence during the Liberian conflict, we would have to prove (or assume) that the Truth and Reconciliation Commission data are representative of the full population of violations. We could begin to measure the representativeness of the TRC data by measuring its overlap with other data sources collected by other organizations in other ways. If all data sources reported similar patterns of violence, and if the overlap between data sources were high (that is,

if many of the same victims were reported to each source), we could cautiously infer that the TRC data represents the population of violations. However, we can neither prove or assume this, because little non-TRC data exists.

A second dataset about violations during the Liberian conflict was commissioned but seemingly abandoned by the United Nations Development Program (UNDP) and later partially analyzed by HRDAG. Conceivably, matching violation records between the TRC data and the UN data might allow analysts to better understand patterns of under-reporting. However, a formal matching project has not been undertaken at this time. While the TRC strove to collect statements from all Liberian counties and from every sector of Liberian society, it is impossible to know the extent to which either the raw data or the relative risk indices represent the true distribution of violence during the Liberian conflict.

Summary of HRDAG Datasets The summary below includes all HRDAG projects, whether active or completed. Much of the data from each of these projects includes specific, identifying information about individual victims. This information is vital to matching records across datasets, which is the key element of the MSE process. However, most data cannot be released publicly with identifying information attached. A few agreements with partner organizations stipulate that data may not be released publicly in any case. While some datasets are posted on the HRDAG website (see http://hrdag.org/resources/data_software.shtml), others are available by contacting HRDAG directly.

[HRDAG data summary table about here]

Conclusion

What do we know about the differential impact of war by sex? More than we used to--data on sex and gender from the Liberian Truth Commission (2009) are significantly better than those from the Truth Commission for El Salvador (1993)--but not very much. In particular, while studies of individual conflicts have improved significantly, analyses of sex, gender and conflict that compare findings across conflicts are still largely impressionistic.

We know that female homicide victims in Kosovo were reported in interviews at slightly lower rates than male homicide victims, but the difference is not statistically significant, and we cannot generalize this finding to all conflicts. We know that women suffered more indirect than direct violence in Timor-Leste, but we do not know whether, or to what extent, this finding can be generalized. Based on all the data surveyed here, we know that more conflict-related violence appears to be directed at men than at women. But again, the extent to which that finding is generalizable, and the magnitude of the gender imbalance, remain unclear. We know that women are more frequently targeted for rape than are men--but we do not know the true extent of the imbalance, or the variation in targeting from conflict to conflict. Nor can we say with certainty that all types of sexual violence are more frequently directed at women.

A research agenda that fully answers these questions likely requires both convenience data of several types and survey data. Women in Kosovo were more likely to be reported as victims of homicides when the information source was an exhumation than when the information source was an interview; women were more likely to report violence in Kosovo when survey enumerators directly questioned randomly sampled heads of household about all types of violence, rather than simply soliciting self-reported, open-ended testimony. Nevertheless, the magnitude of under-reporting for both men and women means that several data sources, including convenience and survey data, are required to craft a reliable, accurate estimate of the pattern and magnitude of violations. To state the case more concretely, 45% of homicides with male victims in Kosovo were reported, versus 40% of homicides with female victims. This 5% differential pales in comparison to the difference between 10,000 estimated total homicides and the 4,400 total reported homicides.

A key recommendation for groups collecting violence data, then, is to find and exploit multiple data sources of multiple types. In addition, where possible, organizations should collect more complete data on non-lethal and indirect violation types, and more complete data on victim sex. However, data collection efforts will never entirely overcome the chaos (and resource pinch) of crisis situations. The unavoidable implication is that questions about patterns and magnitudes of conflict violence require more effort and resources at the “back end” -- analysis -- as well as at the “front end” of data collection. Especially given the advent of automated data-matching and estimation systems, creative assessment and correction of flawed but available data from multiple sources is likely to be the best method for achieving reliable estimates of the gendered impact of armed conflict.

Table 1. Documentation of Male and Female Victims in Raw Kosovo Data

N = 4,400 total documented deaths in raw data, 741 with female victims and 3659 with male victims.

| Dataset | % Males Documented | % Females Documented | Difference (%M-%F) | Significant? (Student's T) |
|----------------------|---------------------------|-----------------------------|---------------------------|-----------------------------------|
| ABA/CEELI Interviews | 37.4 | 36.0 | 1.4 | No |
| HRW Interviews | 42.2 | 40.9 | 1.3 | No |
| Exhumations | 14.3 | 16.2 | -1.8 | No |
| OSCE Interviews | 39.5 | 35.8 | 3.8 | No |

Table 2. Present and Completed HRDAG Projects

| Project Location | Partner Organization | Data Type | # Records | Violation Type(s) | Data on Sex/Gender | Other Information |
|------------------|---|---|---|--|---------------------------------------|---|
| Bangladesh | Human Rights Watch | Interviews; newspaper reports | 367 | Killings | None | |
| Bosnia | Research and Documentation Centre, Sarajevo | Files of 14 governmental and non-governmental organizations | 250,105 | Killings | 83.9% M; 7.3% F; 2.9% "Child"; 5.9% U | See Ball, Tabeau and Verwimp 2007 |
| Chad | Human Rights Watch | 49,000 documents abandoned by Chadian security forces; 792 dossiers from Chadian Assoc. of Victims of Political Repression and Crimes | 12,321 victims; 1,941 unique documents mentioning victims | Detention; Torture; Killings | None | Document flow analysis suggested responsibility of dictator Hissene Habre; see Cruz et al. 2003 |
| Colombia | Coordinación Colombia Europa Estados Unidos (CCEEU) | NGO files | 109 | Killings | 84% M; 12% F; 4% U | |
| | EQUITAS/Families of the Disappeared | Collected records of several organizations (governmental and non-governmental) | ~15,000 | Disappearances (focuses on department of Casanare) | None | See Guzmán et al. 2007 |
| | Policia Nacional | Homicide Database | 73,063 | Killings | 91.8% M; 8.1% F; 0.1% U | See Ball et al. 2007 |
| | Comisión Colombiana de Juristas (CCJ) | Reports to Human Rights NGO | 92,880 | Multiple Violation Types | 58.2% M; 8.4% F; 33.4% U | |
| | Instituto Medicina Legal (IML) | Government Forensic Examiner Records | 131,717 | Lethal Violations | 85.3% M; 13.7% F; 1% U | Includes some non-violent deaths |
| | Indigenous | NGO files | 6,723 | Multiple Violation Types | None | |
| | Vice President's Office on Human Rights | Government Data | 12,936 | Killings, Disappearances | None | |

| Project Location | Partner Organization | Data Type | # Records | Violation Type(s) | Data on Sex/Gender | Other Information |
|------------------|--|--|--------------------------------------|--|--|---|
| | Departamento Administrativo de Seguridad (DAS) | Government Data | 56,077 | Multiple Violation Types | None | |
| El Salvador | Truth Commission for El Salvador | Self-reports to Truth Commission; Reports from other groups | ~7,000 direct reports; ~70,000 total | Killings, Disappearances, Torture, Assault, Rape | None | Data re-scanned from paper annex to Truth Commission Report ca. 2007 |
| | Comisión de Derechos Humanos de El Salvador | NGO files | 22,052 | Multiple Violation Types | None | |
| | El Rescate | NGO files (codings of self-reports to Legal Aid Office of the Archbishop of San Salvador) | 23,100 | Multiple Violation Types | 18% F | |
| India | Ensaaf | Government Human Rights committee; NGO files; Newspaper reports; logbooks of municipal cremation grounds | 21,000 | Disappearances, Extrajudicial Execution | “Overwhelmingly” 18-45-year-old Sikh men | See Silva et al. 2009 |
| Iran | Bouromand Foundation | Web-based database collecting media and self-reports | 12,754 + | Multiple Violation Types | None | Non-conflict human rights violations. Uses HRDAG Analyzer software. See iranrights.org. |
| Kosovo | American Bar Association/ Central and East European Law Initiative | Interviews | 1,755 | Killings | 84% M; 15% F; 1% U | See Ball et al. 2002 |
| | Exhumations | Forensic data | 2,009 | Killings | 84% M; 15% F; 1% U | |
| | Organization for Security and Cooperation in Europe | Interviews | 1,786 | Killings | 81% M; 18% F; 1% U | |
| | Human Rights Watch | Interviews | 685 | Killings | 84% M; 15% F; 1% U | |

| Project Location | Partner Organization | Data Type | # Records | Violation Type(s) | Data on Sex/Gender | Other Information |
|------------------|--|--|--|------------------------------|---------------------------|--|
| | Kosovo data also include migration data from several sources; see Ball 2000, Ball et al. 2002. | | | | | |
| Liberia | Truth and Reconciliation Commission | Self-reports to Truth Commission | 17,169 statements; 163,175 violations | Multiple Violation Types | 47% M; 28% F; 25% U | See Cibelli, Hoover and Krüger 2009 |
| | TRC Diaspora Data | Self-reports to Truth Commission outside Liberia | 6,174 | Multiple Violation Types | 45% M; 33% F; 22% U | |
| | United Nations Development Program | Data contributed to TRC | 21,628 | Multiple Violation Types | 39.7% M; 30.7% F; 7% U | Coded using TRC coding rules. |
| Sierra Leone | Truth and Reconciliation Commission | Self-reports to Truth Commission | 47,418 | Multiple Violation Types | 67.3% M; 32.4% F; 0.3% U | See Conibere et al. 2004 |
| | Coalition for Good Governance/ ABA | NGO files | 25,447 | Multiple Violation Types | 43.9% M; 30.2% F; 25.9% U | |
| | ABA | Household Survey | 64,718 | Multiple Violation Types | 64.2% M; 35.8% F | |
| Timor-Leste | Graveyard Survey | Information from gravestones | 319,000 grave markers | Deaths (Direct and Indirect) | Imputed from first name | See Silva and Ball 2006. Together the three datasets document about 97,000 unique deaths during the CAVR mandate period. |
| | Commission for Reception, Verification and Reconciliation (CAVR) | Self-reports to Truth Commission | 7,668 statements | Multiple Violation Types | 63% M; 37% F | |
| | CAVR | Household Survey | 1,500 households 4,857 total deaths | Deaths (Direct and Indirect) | 55% M; 45% F | |

References

- Amowitz, LL et al. 2002. "Prevalence of War-Related Sexual Violence and Other Human Rights Abuses Among Internally Displaced Persons in Sierra Leone." *Journal of the American Medical Association* 287(4):513-521.
- Baillargeon, S, and L.P. Rivest. 2007. "Rcapture: Loglinear Models for Capture-Recapture in R." *Journal of Statistical Software* 19:5.
- Ball, P. 2001. "Making the Case: The Role of Statistics in Human Rights Reporting." *Statistical Journal of the United Nations ECE* 18:163-173.
- Ball, P et al. 2004. "How Many Peruvians Have Died? An Estimate of the Total Number of Victims Killed or Disappeared in the Armed Internal Conflict between 1980 and 2000." *AAAS Science and Human Rights Program Report*.
- Ball, P., E. Tabeau, and P. Verwimp. 2007. "The Bosnian Book of Dead: Assessment of the Database (Full Report)." *HiCN Research Design Note 5*.
- Ball, Patrick, Tamy Guberek, Daniel Guzmán, Amelia Hoover, and Meghan Lynch. 2007. "Assesing Claims of Declining Lethal Violence in Colombia." Available online at <http://www.hrdag.org/resources/publications/CO-PN-CCJ-match-working-paper.pdf>.
- Ball, Patrick. 2000. *Policy or Panic? The Flight of Ethnic Albanians From Kosovo, March-June 1999*. Washington, DC: American Association for the Advancement of Science.
- Ball, Patrick et al. 2002. "Killings and Refugee Flow in Kosovo, March-June 1999." *Report to the International Criminal Tribunal for the Former Yugoslavia*.
- Carpenter, RC. 2006. "Recognizing Gender-Based Violence Against Civilian Men and Boys in Conflict Situations." *Security Dialogue* 37:1-22.
- Conibere, R et al. 2004. "Statistical Appendix to the Report of the Truth and Reconciliation Commission of Sierra Leone." *HRDAG/ABA/CEELI Report*.
- Cruz, M., et al. 2003. *Preliminary Statistical Analysis of AVCRP and DDS Documents: A Report to Human Rights Watch about Chad Under the Government of Hissene Habre*.
- Guzmán, D et al. 2007. "Missing People in Casanare." *HRDAG White Paper*. Available online at <http://www.hrdag.org/resources/publications/casanare-missing-report.pdf>.

- Hagan, J., and A. Palloni. 2006. "SOCIAL SCIENCE: Death in Darfur." *Science* 313(5793): 1578.
- Hicks, MH, and M Spagat. 2008. "The Dirty War Index: A Public Health and Human Rights Tool for Examining and Monitoring Armed Conflict Outcomes." *PLoS Medicine* 5(12): 1658-1664.
- Hoover, Amelia, et al. 2009. "The 'Dirty War Index' and the Real World of Armed Conflict." Human Rights Data Analysis Group working paper.
- Hoover, Amelia. 2010. "Learning the Hard Way at the ICTY." In *Collective Violence and International Criminal Justice: An Interdisciplinary Approach*, eds. Elie Van Sliedregt and Alette Smeulers. Antwerp: Intersentia.
- Lynch, Meghan, and Amelia Hoover. 2008. "Counting the Uncounted: Multiple Data Systems and the Analysis of Count Data." Presented at the Annual Meetings of the American Political Science Association, Boston.
- Silva, R, and P Ball. 2006. "The Profile of Human Rights Violations in Timor-Leste, 1974-1999." *Report to the Commission for Reception, Verification and Reconciliation of Timor-Leste*. Palo Alto, CA: Benetech.
- Silva, R, et al. 2009. *Violent Deaths and Enforced Disappearances During the Counterinsurgency in Punjab, India: A Preliminary Quantitative Analysis*. Available online at <http://www.hrdag.org/about/india-punjab.shtml>.
- Swiss, S, and P Jennings. 2006. "Documenting the Impact of Conflict on Women Living in Internally Displaced Persons Camps in Sri Lanka: Some Ethical Considerations." *Women's Rights International Report*:1-5.
- Swiss, S. et al. 1998. "Letter From Monrovia-Violence Against Women During the Liberian Civil Conflict." *JAMA-Journal of the American Medical Association-US Edition* 279(8):625-629.
- Taback, N. 2008. "The Dirty War Index: Statistical Issues, Feasibility, and Interpretation." *PLoS Medicine* 5(12):e248.
- Theidon, K. 2007. "Gender in Transition: Common Sense, Women, and War." *Journal of Human Rights* 6(4):453-478.