A Triumph of Remembering: Kosovo Memory Book

By

Professor Michael Spagat
Department of Economics
Royal Holloway, University of London
m.spagat@rhul.ac.uk

December 10, 2014

Abstract

The Kosovo Memory Book Database (KMBD) endeavors to record all deaths due to the war in Kosovo, 1998-2000. The Humanitarian Law Center commissioned the present paper as well as Krüger and Ball (2014) to evaluate the quality of this database. My work takes advantage of the serendipitous existence of two separate and independent sets of statistical estimates of the number of war deaths in Kosovo. These two projects used entirely different data and statistical methodologies so it is ambitious to hope for consistency between the two of them, let alone between these two sets of estimates and the death counts of KMBD. Nevertheless, I find excellent consistency across the three projects for overall totals, breakdowns by gender and age ranges and patterns over time and space. These findings suggest that the database has reached a near-saturation point. The fact that Krüger and Ball (2014) searched extensively over lists of war victims and were unable to identify a new death that clearly belongs in KMBD reinforces this conclusion, as does the rigorous methodology of the database itself. I also show that lists of possible fatalities that could potentially be added to KMBD if more positive evidence becomes available to support their inclusion could be absorbed into KMBD without substantially changing the patterns of deaths over time or space. I conclude that KMBD is a very high-quality database that is likely to capture virtually all deaths in the Kosovo war.

1 This is one of two evaluations of the Kosovo Memory Book database commissioned by the Humanitarian Law Center. The other one is written by Dr. Jule Krüger and Dr. Patrick Ball of HRDAG.

2 I am hugely grateful to Figo Sze-Yeung Lau of Oxford University. We worked closely together analyzing earlier public versions of the Kosovo Memory Book database and jointly developed most of the insights I present in this report. Unfortunately, I had to leave him behind for this report due to a confidentiality agreement over the latest version of the database which is not yet in the public domain. I also thank Patrick Ball and Jule Krüger for their insightful comments on a draft of this report.
1. Introduction

The charter for the recognition of every casualty of armed violence, launched on September 15, 2011 (Every Casualty, 2011a), calls upon states and other relevant armed conflict actors to ensure that all conflict casualties are promptly recorded, correctly identified and publicly acknowledged. Appropriately, the Humanitarian Law Center (HLC) and the Humanitarian Law Center – Kosovo (HLCK) played a central role in the launch (Every Casualty, 2011b). Indeed, HLC and HLCK began work on their Kosovo Memory Book project long before the Every Casualty project was conceived and it is hard to think of any project in the world that is more in tune with the Every Casualty goals than is the Kosovo Memory Book (Humanitarian Law Center and Humanitarian Law Center-Kosovo, 2014a).

The core goals of both Every Casualty and Kosovo Memory Book are to record, memorialize, remember and acknowledge every single conflict death – individual by individual. These groups view conflict recording as an essential human activity, akin to burial rituals.

In this report I study statistical estimates and counts of fatalities in the Kosovo war. This analysis may seem far removed from the core purpose of Kosovo Memory Book which is to remember the human victims of the war one by one. Yet I suggest that the material presented in this paper does, indeed, serve a core purpose of the Kosovo Memory Book because it supports the case that HLC and HLCK have built an accurate and virtually comprehensive record of fatalities due to the Kosovo war.3 This is important because HLC and HLCK can only achieve their core goals if they accurately list virtually every single victim of the war. Thus, the present report provides good evidence that Kosovo Memory Book is an exemplary success within the framework established by the Every Casualty project.

The structure of this report is as follows. In sections 2 and 3 I compare the violent death counts of the Kosovo Memory Book database with those of two

---

3 There will always be the possibility that compelling evidence of a new death will surface so no one should every claim that any such database is fully comprehensive.

4 Krüger and Ball (2014) address the same questions from many different angles and come to the similar conclusions.
different statistical estimates. I find strong consistency across the sources on the total number of deaths as well as on how these deaths are distributed across time, space and demographic categories. In section 4 I draw attention to the work of Krüger and Ball (2014) searching many lists of Kosovo war victims, including some that HLC and HLCK have not had access to. They do not find any deaths that clearly belong in the Kosovo Memory Book database but that are not already there. The rigorous methodology of the database renders this finding understandable (Humanitarian Law Center and Humanitarian Law Center-Kosovo, 2014b). Section five considers candidates currently held to the side for potential integration into the database at a later stage if compelling evidence surfaces that they belong in there. I show that even transfer of a large proportion of these potential deaths into the main database would have no discernable effect on the Kosovo Memory Book database as a whole. I draw some conclusions in section six.

2. Kosovo Memory Book versus the Spiegel and Salama Estimates

a. Why this comparison is interesting and important to us

The main goal of Spiegel and Salama (2000, hereafter “S&S”) was to estimate the number of people killed in the Kosovo war. A secondary goal was to give relatively detailed breakdowns on how estimated deaths were spread across time and demographic groups. In particular, S&S gave monthly estimates and also disaggregated their totals into males and females belonging to three broad age classes - six demographic groups in total. This level of detail enables rich comparisons between S&S and the Kosovo Memory Book database (KMBD).

S&S based their estimates on data gathered by a sample survey.⁵ This means they started by drawing a random sample of households in Kosovo. They then conducted a household interview to record, among other things, the number of violent deaths suffered by household members. The interview also recorded the presence of other live and dead household members, accounting for births, non-violent deaths, in-migration and out-migration.

---

⁵ Spagat (2012) gives an overview of the use of sample surveys for measuring deaths in armed conflict.
The key step in this approach is to project the violent death rate found within
the sampled households onto the entire population of Kosovo. The estimate of
violent deaths arrived at by this technique will be correct on average under certain
ideal conditions that the survey strived to attain. This averaging is over all the
possible samples that could have been drawn via the procedures S&S used to draw
the one sample that they analyzed. There are, in fact, reasons to believe that more
often than not survey estimates are below the true number they are trying to
estimate even though they get the right answer on average (Spagat, 2009, slide 9).

A survey departs from ideal conditions if, for example, the household interview
does not accurately elicit the correct information or the sampling procedures
systematically favor selecting households that have experienced more (or less)
violence than is typical for the population as a whole. No survey, including S&S,
satisfies all the ideal conditions but S&S does appear to be a fairly high-quality
survey so we would expect it to produce reasonable estimates.

S&S follow the standard practice of using their sample to quantify how far
their estimate may deviate from the true number of deaths, always operating under
the maintained assumption that the ideal conditions for a survey are satisfied. This
calculation aims to capture the characteristics of a plausible range of household
samples that could have been randomly selected; each sample will lead to a different
estimate of the number of violent deaths. The resulting range, known as a “95 %
certainty interval,” incorporates uncertainty due to the sample being random but
does not normally account for possible departures from ideal conditions.

The above discussion applies to sub-categories of deaths as well as to total
deaths. However, we must be aware that the more specific the category of death the
more uncertain the survey-based estimates will be. Estimates of violent deaths in a
month will be less reliable than estimates of violent deaths in a year. Estimates for

---

6 This asymmetry results from the fact that households without violent deaths greatly outnumber
households with violent deaths. This means that most random sample will contain too many
households without violent deaths while a fair number of the samples that contain too many violent
deaths will actually contain far too many violent deaths.
males will be less reliable than estimates for females plus males. Nevertheless, under ideal conditions these estimates should be correct on average.

The S&S survey appears to be a high-quality one. Thus, the above discussion gives us a valid theoretical basis to believe that the overall S&S estimate is likely to be close to the true numbers of violent deaths in the Kosovo war. Estimates for more specific categories such as months, genders or age ranges are more subject to random variation than the overall estimate is but can still be expected to be reasonably close to correct values. Confidence intervals, which quantify the uncertainty of working with a random sample rather than the entire population of Kosovo, provide at least a partial handle on the uncertainty surrounding these estimates.

If the KMBD database is a virtually comprehensive list of deaths in the Kosovo war then the total number of violent deaths in the KMBD database should be similar to the S&S estimate for violent deaths. Moreover, KMBD totals by month, gender and age should also resemble the S&S estimates. We must remember that there is a large random component within the S&S estimates, especially in sub-categories, so we should not be surprised to see a few strong deviations between KMBD and S&S numbers. Still, they should be quite close to one another on average.

b. The comparisons

We eliminate the 281 “deaths caused by war” from KMBD to put it on a comparable basis with the S&S estimates of “war-related trauma” which do not include non-violent deaths that can be attributed to the war as KMBD does. We also limit KMBD to just the period January 1, 1998 through September 30, 1999 since this is the period covered by S&S. We are left with 12,965 violent deaths in KMBD that are fit for comparison with S&S.

The totals for the two projects for the whole S&S period are close – 12,965 for KMBD versus 12,010 with a 95% confidence interval of 5,500 to 18,300 for S&S. As mentioned above, survey estimates under ideal conditions should come out below the true number more often than not so KMBD and S&S line up very much like one
would expect if both projects are accomplishing their goals, specifically, recording every war death for KMBD and estimating the true number of war deaths for S&S.

Table 1. KMBD Versus S&S – Violent Deaths by Age and Gender

<table>
<thead>
<tr>
<th>Age Range/Gender</th>
<th>Spiegel and Salama Estimates</th>
<th>Kosovo Memory Book</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower Limit</td>
<td>Central Estimate</td>
</tr>
<tr>
<td>0-14 Years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>0</td>
<td>161 (44%)</td>
</tr>
<tr>
<td>Females</td>
<td>0</td>
<td>201 (56%)</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>362 (100%)</td>
</tr>
<tr>
<td>15-49 Years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>1694</td>
<td>5421 (91%)</td>
</tr>
<tr>
<td>Females</td>
<td>0</td>
<td>510 (9%)</td>
</tr>
<tr>
<td>Total</td>
<td>1694</td>
<td>5931 (100%)</td>
</tr>
<tr>
<td>50+ Years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>2164</td>
<td>5176 (91%)</td>
</tr>
<tr>
<td>Females</td>
<td>0</td>
<td>541 (9%)</td>
</tr>
<tr>
<td>Total</td>
<td>2164</td>
<td>5717 (100%)</td>
</tr>
<tr>
<td>All Ages</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>3858</td>
<td>10758 (90%)</td>
</tr>
<tr>
<td>Females</td>
<td>0</td>
<td>1252 (10%)</td>
</tr>
<tr>
<td>Grand Total</td>
<td>5500</td>
<td>12010 (100%)</td>
</tr>
</tbody>
</table>

Lower and upper limits refer to 95% confidence intervals. Calculations for S&S are based on their table 2, except for the confidence interval for the "Grand Total" which is their overall range.

There is another reason why it is notable that the KMBD total is higher than the S&S central estimate of 12,010. One might expect the KMBD count to be lower
than the true number of violent deaths since it is extremely difficult to detect every single death on a case-by-case basis.\(^7\) Therefore, exceeding a valid statistical estimate like the S&S one is a sign the KMBD may be virtually comprehensive.

The S&S paper gives a breakdown by gender and age categories, enabling a direct comparison with KMBD along these dimensions as shown in table 1. These detailed comparisons are consistent with the notion that KMBD is virtually comprehensive. KMBD counts exceed the S&S central estimates for all sub-categories except males aged 50+ and are always well below the upper limit of the 95% confidence interval for S&S. It seems clear that 50+ males were disproportionately targeted during the war although the S&S estimate appears to exaggerate the extent to which this is true, most likely due to a random variation within the sample.

Figure 1 compares monthly time series for the KMBD and S&S.\(^8\) The two series track each other remarkably well, particularly considering that the S&S series is based on only 50 clusters of 24 households and is, therefore, subject to rather strong random variation once the estimates are disaggregated down to a monthly level. Unfortunately, we cannot quantify this randomness with confidence intervals because the original S&S data have been lost.\(^9\)

\(^7\) It is, however, possible for a documented count like KMBD’s to exceed the true number of deaths if there are duplicates or false deaths lying undetected within the database.
\(^8\) Figo Sze-Yeung Lau of Oxford University and I built the monthly numbers for the S&S time series via painstaking inspection of figure 2 of the S&S paper.
\(^9\) Personal correspondence with Paul Spiegel.
The correlation between the two monthly series is 0.86 on a scale from 0 to 1 where 1 would indicate perfect tracking and 0 would indicate a complete absence of tracking. Even if KMBD is a perfect record of all violent deaths in the Kosovo war and S&S is an exceptionally high-quality survey we would still never expect the correlation to be 1.0 since the S&S estimate comes from a random sample. Based on the available information 0.86 appears to be an excellent result.

In summary, the KMBD-S&S comparisons come out quite well. They are at least consistent with KMBD being a virtually complete and very high quality database.

3. Kosovo Memory Book versus the Ball et al. Estimates
   
a. Why this comparison is interesting and important to us
Ball, Betts, Scheuren, Dudukovic and Asher (2002) (hereafter BBSDA) made statistical estimates of violent deaths in the Kosovo war that were entirely independent of S&S and used a completely different method known as “multiple systems estimation (MSE).”

It is interesting to compare the BBSDA estimates with KMBD numbers for the same reason it is interesting to compare S&S with KMBD. Under ideal conditions the BBSDA method should give the true number of violent deaths on average so if total deaths recorded by KMBD are close to this true number then KMBD numbers should also be close to BBSDA estimates subject to the caveat that the BBSDA estimates have random noise built into them.

There are three main reasons why this second round of comparisons adds valuable new information to our analysis. First, S&S is based on a random sample so its sample could turn out to be rather unrepresentative of the population even though it was a well conducted survey. If the S&S estimate does suffer from such a random aberration we would have a decent chance of catching this problem by looking at a second estimate. Second, BBSDA use a different statistical technique than S&S do so if one of these projects is inaccurate due to strong departures from ideal conditions for the application of the technique (rather than just random variation) then there is a reasonable chance that this problem will be detected by the other project. Third, we can compare KMBD and BBSDA by regions, something that was not possible with S&S since this paper does not provide geographical information.

BBSDA’s statistical technique is, unfortunately, more complicated, less intuitive and less familiar than are S&S’s survey methods. The survey approach works by controlling and understanding the random procedures by which data are gathered. Within the context of estimating war deaths MSE is normally invoked when data have already been gathered by multiple organizations that have not followed the controlled randomization procedures that underpin sample surveys. Researchers

10 Manrique-Vallier, Price and Gohdes (2013) and Jewell, Spagat and Jewell (2013) provide overviews of the application of MSE techniques to estimating deaths in armed conflict.
then try to model the processes whereby deaths are captured, or not captured, by these multiple data-gathering operations. The key driver of model construction is how recorded deaths across the multiple recording systems overlap with one another. In other words, the analysis focuses on which of the multiple sources capture, and which fail to capture, each known death. Intuitively, if every known death appears in almost all recording systems then it is likely that, between all the systems, almost all deaths have already been captured. At the other extreme, if most deaths appear in only a single system then it is likely that many deaths have actually occurred but were not captured by any of the systems.

If there are false or inaccurate records within underlying sources used for MSE then there is a leveraged inflation of the estimates, i.e., X false records will add more than X deaths to the estimate. The BBSDA estimates were based on lists of reported war dead that were available shortly after the war finished so it is likely that some of these reported deaths would not have withstood the sustained scrutiny of HLC and HLCK over the last decade. Indeed, over the course of its work KMBD has found 3,258 reported deaths that do not meet its standards for inclusion in the database (Humanitarian Law Center and Humanitarian Law Center-Kosovo, 2014b). Some of these “not war victims” may have been incorporated into the BBSDA estimate. Thus, it would not be surprising if the BBSDA estimate was a bit too high.

b. The comparisons

The BBSDA estimates cover the period from March 20, 1999 through June 20, 1999. Although this is a short period it covers a strong majority of all the deaths. For the sake of these comparisons we restrict KMBD to violent deaths just as we did with the KMBD-S&S comparisons.

The BBSDA overall estimate is 10,356 with a 95% confidence interval of 9,002 to 12,122. The KMBD number for this period is 9,790, placing it right in the middle of the BBSDA range. As noted at the end of section 3a, we might have predicted the BBSDA estimate to come out a bit high due to the likely presence of false or inaccurate reports within the sources feeding into that estimate.
Table 2 gives BBSDA and KMBD numbers broken down by region. Again, the two datasets are fairly consistent with each other although there is more random variation at this lower level of aggregation.

Table 2. Regional comparisons: KMBD versus BBSDA

<table>
<thead>
<tr>
<th>Project/Region</th>
<th>North</th>
<th>South</th>
<th>East</th>
<th>West</th>
</tr>
</thead>
<tbody>
<tr>
<td>KMBD</td>
<td>3138</td>
<td>1171</td>
<td>1501</td>
<td>3662</td>
</tr>
<tr>
<td>BBSDA</td>
<td>3925</td>
<td>1606</td>
<td>1827</td>
<td>3188</td>
</tr>
</tbody>
</table>

Region is unknown or outside Kosovo for a small number of KMBD deaths so the KMBD numbers do not quite add up to 9790.

Finally, figures 2-5 compare time series, using two-day estimates from BBSDA. Two-day estimates will have substantial randomness mixed into them so they can deviate considerably from the KMBD numbers without necessarily signaling a serious problem. I try to reduce the impact of the randomness by presenting four-day, six-day and eight-day windows. It is clear that widening the windows irons out many of the fluctuations in the BBSDA estimates. Also the correlations between the KMBD and BBSDA numbers tend to rise with window size. These time series comparisons are not as visually striking as the monthly comparison between KMBD and S&S. However, I argue that they are similarly compelling, given the challenge of matching up such disaggregated numbers.

It is worth noting that BBSDA numbers do tend to be systematically higher than the KMBD ones at the beginning of the period and systematically lower at the end of the period with these differences showing themselves more and more clearly as we move to wider and wider time windows. These differences are persistent enough over time that they are unlikely to be random variations. They are much more likely to be explained by problems in one or more of the sources underlying

---

11 The numbers for the wider windows are always obtained by adding up the numbers from the two-day windows.
Figure 2. Kosovo Memory Book Versus Ball et al.  
March 20, 1999 through June 20, 1999

Figure 3. Kosovo Memory Book Versus Ball et al.  
March 20, 1999 through June 20, 1999
Figure 4. Kosovo Memory Book Versus Ball et al.
March 20, 1999 through June 20, 1999

Figure 5. Kosovo Memory Book Versus Ball et al.
March 20, 1999 through June 20, 1999
the BBSDA estimates than by a problem with the KMBD numbers themselves.

To summarize, the KMBD versus BBSDA comparisons also suggests a high level of quality and completeness for the KMBD database. It appears that KMBD passes every test it faces.

Some readers of this document may assume that it is typical for multiple estimates and counts of war deaths to be extremely consistent with one another. It is not.\textsuperscript{12} Such consistency is rare and unexpected, reflecting well on all three projects. The results of the above comparisons should give us confidence in Kosovo Memory Book which is a truly remarkable project.

4. A Third Argument Suggesting that Kosovo Memory Book is Virtually Comprehensive

Section 6.3 of the companion report to the present paper (Krüger and Ball, 2014) provides another reason to believe that KMBD is virtually comprehensive. Krüger and Ball (2014) evaluate all available sources recording deaths in the Kosovo war and are unable to find deaths that clearly belong in KMBD and are not already there. They do find a few candidates but these cases lack enough detail to make a strong case to include them in the database.

The fact that KMBD did not have access to four of the sources analyzed by Krüger and Ball (2014) adds further weight to this finding. It would be striking if even one unprocessed source was considered but had no impact on KMBD. That this happened four times is impressive and suggests that KMBD is near a saturation point in terms of recording war deaths.

The KMBD methodological description highlights the systematic and sustained effort that has been applied to the construction of KMBD Humanitarian

\textsuperscript{12} For example, there are rather dramatic inconsistencies between various survey estimates of violent deaths in the Iraq conflict. There is no up to date survey at the moment but Spagat (2010) and Spagat and Dougherty (2010) discuss many of the problems in this literature.
Law Center and Humanitarian Law Center-Kosovo, 2014b). This document underscores the rigor of work and supports the contention that the database in a virtually complete record of war deaths in Kosovo.

5. Can the Potential Victims List Substantially Change the War Victims List?

Figure 6 gives monthly time series for KMBD war victims and for KMBD war victims plus potential victims. It is clear from the picture that the shape of the KMBD fatality numbers would not change substantially even if many of the potential victims are eventually shifted onto the actual victim list.\(^{13}\)

The main reason the curves in figure 6 track each other so well is because there are many more actual victims than potential victims (13,517 versus 1,603) so the curve for actual plus potential victims greatly resembles the curves for actual deaths alone. The time series for potential victims is only somewhat correlated with the time series for actual war victims (Table 3). Potential victims and war victims are also weakly correlated at the regional level (correlation of 0.13).

As a matter of principle, KMBD aspires to record every single victim of the war. So the possibility of shifts from potential to actual victims is important for the project. Nevertheless, from the perspective of a researcher using the KMBD data these shifts could hardly be important since it is unlikely that any research result could hinge on whether or not potential victims are added into the main database.

\(^{13}\) The separation in the two curves for January 1999 is a coding anomaly since all potential deaths for 1999 with unknown months are coded for January of 1999.
Table 3. Correlations between War Victims, Potential Victims and War Victims plus Potential Victims at the Daily and Monthly Levels

<table>
<thead>
<tr>
<th>Correlations</th>
<th>War victims</th>
<th>Potential Victims</th>
<th>War Victims plus Potential Victims</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>daily</td>
<td>monthly</td>
<td></td>
</tr>
<tr>
<td>War Victims</td>
<td>1.0</td>
<td>1.0</td>
<td>0.06</td>
</tr>
<tr>
<td>Potential Victims</td>
<td>0.06</td>
<td>0.25</td>
<td>1.0</td>
</tr>
<tr>
<td>War Victims plus</td>
<td>0.90</td>
<td>0.99</td>
<td>0.49</td>
</tr>
<tr>
<td>Potential Victims</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6. Conclusion

There is excellent consistency between KMBD, S&S and BBSDA ranging across time, space, demographics and aggregate totals. Krüger and Ball (2014) tried but could not identify an appropriate death to add to KMBD. The general shape of the KMBD time series would not change substantially even if more evidence becomes available in the future and a number of potential deaths are transferred from the list of potential deaths onto the list of actual deaths. In short, KMBD passes all the tests I subject it to.

Yet the present paper is positioned within a larger picture of quality and consistency, much of which is documented in Krüger and Ball (2014) and displayed in Humanitarian Law Center and Humanitarian Law Center-Kosovo (2014b). Together these three papers provide ample evidence of high quality and near comprehensiveness of the KMB database.

It is very unusual for a project documenting war deaths on a case by case basis to progress to a point where it captures virtually every single one of them.14 KMBD appears to have reached this point while simultaneously providing quite a wide range of accurate information on each victim.

Kosovo Memory Book is an extraordinary achievement that stands as a monument both to its victims and to the human spirit.

---

14 Sutton (1994) for Northern Ireland and B’Tselem (2014) for the Israeli-Palestinian conflict are likely to be virtually complete enumerations of fatalities in those conflicts. The Bosnian Book of the Dead is also likely to come close to a complete enumeration (Ball, Tabeau and Verwimp, 2007)
Bibliography


Humanitarian Law Center and Humanitarian Law Center-Kosovo (2014b) The database methodology of the Kosovo Memory Book.


