Estimating Undocumented Homicides with Two Lists and List Dependence

Kristian Lum & Patrick Ball

April 2, 2015

Human Rights Data Analysis Group
everybody counts.

1 Introduction

Homicides tend to be hidden from public view. Perpetrators are often motivated to conceal the crime, and victims' families may be afraid to denounce the violence; concealment and the families' fear may be most acute when the perpetrators of the crime are state authorities, like the police. Consequently, lists of homicides tend to be partial, and they tend to emphasize victims with high social visibility: victims who are relatively well-known, and whose killing occurs in daylight, in urban areas, and in view of bystanders motivated to report the crime. Other killings without these aspects more frequently remain hidden from public knowledge.

When two or more groups provide lists of victims of homicide, it is possible to estimate the total population of victims, including those who were not documented on any of the lists being used. The technique is called *capture-recapture* or *multiple systems estimation* (MSE; an introduction to the method is here). Intuitively, the more overlaps among the lists, the more plausible it is that the population they are drawing from is small.

A standard assumption in MSE is that the lists are *statistically independent* (see Q13, here), i.e. that an incident being recorded on one list makes the incident no more or less likely to be recorded on the other lists. Another standard assumption is *homogeneity of recording probabilities* (see Q11, here), i.e. that the probability of recording patterns does not vary across the population being estimated. Heterogeneity in recording probabilities can induce list dependence (International Working Group for Disease Monitoring and Forecasting, 1995). Thus, the independence assumption is rarely true, but with three or more lists we can estimate the dependence among subsets of the lists, following the method proposed by Bishop et al. (1975). In estimates made by the Human Rights Data Analysis Group (HRDAG), we use three or more lists to take advantage of the additional lists to estimate the list dependence rates.

Whether estimates produced under the assumption of independence over- or underestimate the true population size depends on correlations between inclusion in different lists. Lists are positively correlated if the appearance of an incident on one list makes it more likely that the incident appears on the other list. One mechanism through which this can occur is if the both lists are compiled using some of the same underlying data sources. Positive list correlation also occurs between lists collected by projects that share similar social constituencies, for example, lists of victims collected by police and by municipal social workers may tend to draw from communities that trust the government, while communities that do not trust the government may avoid both police and government social service projects. Lists are negatively correlated if the appearance of an incident on one list makes it less likely that the incident was recorded on the other list. This can occur if the groups gathering data tend to focus their efforts on different geographic regions or periods of time, or if one documentation group draws from one political party while another documentation group draws from a competing party. When dependence between two lists is positive, the two-list independence estimator will be biased *downward*, and when list dependence is negative, this estimate will be biased *upward*. In practice, we have found that most list dependence is positive.

In this document, we propose a method to include a correction for list dependence in the two list case, producing a range of estimates. In essence, we propose performing a sensitivity analysis to the independence assumption. For the values we use in the correction, we derive list dependence measures from contextually similar projects where we believe the underlying data generating processes was similar to the data collection done by the two groups recording from our target population. That is, we select other data sets that we believe exhibit similar list dependence properties to the two lists of records from our target population. Thus, we can estimate a population total using only two lists that accounts for list dependence. This approach assumes that the list dependence in the two lists from our target population is comparable to the list dependence in other projects (and populations) where we have three or more lists.

As an example of our approach, we derive measures of list dependence from previous work we have done studying homicides in Colombia (Guzmán et al., 2011), Kosovo (Ball et al., 2002), Guatemala (Ball, 1999), Sierra Leone, and Syria. We use the list dependence values estimated in these countries to produce a range of estimates of the number of homicides committed by police in the United States using only two lists. The multiple systems estimation approach is introduced with two lists in Section 2. In Section 2.1, we briefly review the log-linear approach to multiple systems estimation and its properties. Our sensitivity analysis methodology is outlined in Section 4. In Section 5, we provide an example for our methodology, using data on homicides from other countries to perform a sensitivity analysis on the number of US police homicides reported in a recent report by the Bureau of Justice Statistics (hereafter the "BJS report"). We noted previously that the BJS estimates were likely biased downward. Our findings here provide a range of estimates of police homicides after accounting for the effect of list dependence between the sources used by the BJS. Our approach assumes that the estimated range of list dependences from the other projects is relevant to the US police homicides data. This serves as a sensitivity analysis to the BJS report that relaxes their assumption of list independence and explores the likely consequences of this assumption.

2 Estimating Total Population Size Using Two Lists

Consider a setting in which two partial lists of incidents – e.g. homicides by police – have been gathered. For simplicity, we assume that every incident is uniquely identifiable, for example through the social security numbers of the victims. This results in data consisting of three counts: the number of incidents appearing only on the first list (n_{10}), the number of incidents that appear only on the second list (n_{01}), and the number that appear on both lists (n_{11}). The goal of statistical inference in this setting is to estimate the number incidents that did not appear on either of the lists, n_{00}, or equivalently, the total number of incidents, $N = n_{00} + n_{10} + n_{01} + n_{11}$.

As described previously, estimation of N with only two lists requires the assumption of list independence. Under independence, a common estimate of the total population size can be derived in the following way. Suppose that the vector of counts $\boldsymbol{n} = (n_{00}, n_{01}, n_{10}, n_{11})$ has likelihood

$$\boldsymbol{n}\big|N \sim \text{Multinomial}(N, \boldsymbol{p}) \tag{1}$$

where $\boldsymbol{p} = (p_{00}, p_{01}, p_{10}, p_{11})$ satisfies $\sum_{\boldsymbol{i}\in\{0,1\}^2} p_{\boldsymbol{i}} = 1$. When N is known, the maximum likelihood estimate (MLE) for $\boldsymbol{p}$ is $\hat{p}_{\boldsymbol{i}} = n_{\boldsymbol{i}}/N$. Thus, the MLE for the marginal probability of being recorded on the first list is

$$\hat{p}_{1+} = \hat{p}_{10} + \hat{p}_{11} = \frac{n_{10} + n_{11}}{N} = \frac{n_{1+}}{N},$$

and similarly, $\hat{p}_{+1} = n_{+1}/N$, where the subscript $+$ indicates summation over the corresponding index. Similarly, for known N, the MLE of the probability of an incident appearing on both lists is $\hat{p}_{11} = n_{11}/N$. Under independence, it must also be true that the probability of appearing on both lists is $p_{+1}p_{1+}$, so

$$\tilde{p}_{11} = \hat{p}_{+1}\hat{p}_{1+} = \frac{n_{+1}}{N}\frac{n_{1+}}{N}$$

is another estimate of p_{11}. Equating these two estimates gives

$$\tilde{p}_{11} = \frac{n_{1+}}{N}\frac{n_{+1}}{N} = \frac{n_{11}}{N} = \hat{p}_{11}, \tag{2}$$

resulting in $\hat{N} = (n_{1+}n_{+1})/n_{11}$, an estimate under multinomial sampling of the population size under independence. With only two lists, the independence assumption is necessary to identify the model, as it is not possible to estimate both the total population size *and* the joint probability of capture on both lists.

2.1 Log-Linear Representation

A more easily extensible and computationally convenient approach to two-list population estimation replaces the Multinomial likelihood with the Poisson in the log-linear model framework (Lang (1996), Cormack and Jupp (1991), Cormack (1992)). In this framework, the entries of $\boldsymbol{n}$ are assigned a Poisson likelihood with parameter $\mu_{\boldsymbol{i}}$, where

$$\mu_{\boldsymbol{i}} = e^{\beta_0 + \beta_1 i_1 + \beta_2 i_2}, \tag{3}$$

where $\boldsymbol{i} \in \{0,1\}^2$, so that $n_{\boldsymbol{i}}$ is the number of incidents appearing on the lists for which i_j is 1 but not on the lists for which $i_j = 0$ for $j = 1, 2$. The number of unrecorded homicides, n_{00}, is estimated by its expectation, $\mu_{00} = e^{\beta_0}$. Conditioning on the population total N, the Poisson model implies a multinomial model, (1), with $p_{\boldsymbol{i}} = \frac{\mu_{\boldsymbol{i}}}{\sum_{\boldsymbol{i}} \mu_{\boldsymbol{i}}}$.

Let us consider the interpretation of the β parameters in the above model. Consider $\mathbb{E}[n_{1+}]/\mathbb{E}[n_{0+}] = \mu_{1+}/\mu_{0+} = p_{1+}/p_{0+}$, the odds ratio of appearing on the first list. Substituting the expressions for these quantities under equation 3 gives

$$\frac{\mu_{10} + \mu_{11}}{\mu_{00} + \mu_{01}} = \frac{e^{\beta_0+\beta_1} + e^{\beta_0+\beta_1+\beta_2}}{e^{\beta_0} + e^{\beta_0+\beta_2}} = e^{\beta_1}.$$

Similarly, β_2 represents the log odds of capture on the second list.

Allowing a nonzero interaction term between the two inclusion indicators, we have the following model:

$$\mu_{\boldsymbol{i}} = e^{\beta_0+\beta_1 i_1+\beta_2 i_2+\beta_{12} i_1 i_2}. \tag{4}$$

The pairwise interaction term, β_{12}, represents the log odds ratio:

$$\frac{\mu_{11}/\mu_{01}}{\mu_{10}/\mu_{00}} = \frac{e^{\beta_0+\beta_1+\beta_2+\beta_{12}}/e^{\beta_0+\beta_2}}{e^{\beta_0+\beta_1}/e^{\beta_0}} = e^{\beta_{12}}. \tag{5}$$

If $\beta_{12} = 0$, then the odds ratio equals one, which is equivalent to the inclusion on lists 1 and 2 (i_1, i_2) being independent.

Unfortunately, this model is not identified, as we only have three data points, n_{01}, n_{10}, and n_{11}, with which to estimate four parameters, β_0, β_1, β_2, and β_{12}. If, however, we somehow knew or could estimate β_{12} (or had a range of plausible values) from other data sources, this information could easily be incorporated into the log-linear model by fixing β_{12} at this value (or each value in its range) and estimating the other parameters.

3 Population Estimation with more than two lists

Suppose that in addition to the two lists discussed above, we now receive a third list to be used in our analysis. This can easily be incorporated in the log-linear framework, as adding more lists to the model simply requires adding more terms to the log-linear representation. With more than two lists, the pairwise interaction terms become identified. To extend the notation to a general case of J partial lists of incidents, let $n_{\boldsymbol{i}}$ be the count of the number of incidents appearing on the lists in the set $\{j : i_j = 1\}$ but not the lists in the set $\{j : i_j = 0\}$ for $\boldsymbol{i} = [i_1...i_J]$, and $j \in 1, \ldots, J$. For example, analogous to the two list setting, n_{1011}, denotes the number of incidents recorded on lists one, three, and four, but not on list two; n_{1+11} denotes the number of elements recorded on lists one, three, and four. In this setting, there is no simple, closed-form estimate for the total population size as in the two list case, and estimation is typically done via maximum likelihood in a Poisson log-linear model, i.e.

$$\mathbb{E}[n_{\boldsymbol{i}} \mid \theta] = \mu_{\boldsymbol{i}} = \exp\left\{\theta_0 + \sum_{j}^{J} \theta_t i_t + \sum_{j=1}^{J} \sum_{k>j} \theta_{jk} i_j i_k\right\}. \tag{6}$$

As in the two list case, a model that includes only the main effects (θ_t) corresponds to an independence model. List dependence is induced by including the θ_{jk} pairwise interaction terms. Unlike in the two list case, here there is sufficient data to estimate the pairwise interactions when $n_{00\ldots0}$ is unobserved. From J lists, $2^J - 1$ list overlap counts, $n_{\boldsymbol{i}}$, can be calculated. A model with J lists that includes an intercept, main effects, and all pairwise interactions contains $1 + J + J(J-1)/2$ parameters. For $J > 2$, $2^J - 1 > 1 + J + J(J-1)/2$. Thus unlike the two list case, in the three or more list case, the pairwise interaction terms are estimable.

4 Sensitivity Analysis

Suppose we are trying to estimate population A with list intersection counts $\{n_{01}, n_{10}, n_{11}\}$, from which we have only two partial lists. As described above, we do not have sufficient data to include an interaction term and fit model 4. Instead, we will use a plausible range of values for β_{12} derived from a collection of other datasets that analyze other populations. This collection consists of datasets with more than two lists, and analyzed using versions of model 6, i.e. log-linear models with interaction terms. We produce a range of plausible population size estimates for population A via the following procedure:

1. For each population d in the collection of datasets, let J_d be the number of lists of recordings from that population. (We chose our collection such that $J_d > 2$.)

2. Fit model 6 to each population d.

3. Consider all $\binom{J_d}{2}$ pairs of lists. For each pair r, compute the marginal log odds ratio using $\hat{\mu}_{\boldsymbol{i}}$ estimated from model 6, e.g. for $r = \{2, 3\}$:

$$\tilde{\beta}_{12}^{(d,r)} = \log \frac{\hat{\mu}_{+11+\ldots+} / \hat{\mu}_{+01+\ldots+}}{\hat{\mu}_{+10+\ldots+} / \hat{\mu}_{+00+\ldots+}} = \log \frac{\dfrac{\sum_{\boldsymbol{i} \in \{0,1\}\times\{1\}\times\{1\}\times\{0,1\}^{J_d-3}} \hat{\mu}_{\boldsymbol{i}}}{\sum_{\boldsymbol{i} \in \{0,1\}\times\{0\}\times\{1\}\times\{0,1\}^{J_d-3}} \hat{\mu}_{\boldsymbol{i}}}}{\dfrac{\sum_{\boldsymbol{i} \in \{0,1\}\times\{1\}\times\{0\}\times\{0,1\}^{J_d-3}} \hat{\mu}_{\boldsymbol{i}}}{\sum_{\boldsymbol{i} \in \{0,1\}\times\{0\}\times\{0\}\times\{0,1\}^{J_d-3}} \hat{\mu}_{\boldsymbol{i}}}}. \tag{7}$$

4. For each population d and pair r, fit the following Poisson regression model: $\log E[n_{\boldsymbol{i}} \mid \beta] = \beta_0 + \beta_1 i_1 + \beta_2 i_2 + \tilde{\beta}_{12}^{(d,r)} i_1 i_2$. Output the resulting population size estimate $\hat{N}^{(d,r)} = \sum_{\boldsymbol{i}} n_{\boldsymbol{i}} + e^{\hat{\beta}_0}$.

Our estimate of the marginal log odds ratio, $\tilde{\beta}_{12}$, will be zero if the model used to estimate the $\mu_{\boldsymbol{i}}$ were the independence model. Thus, to estimate the marginal log odds ratio, it was necessary to have more than two lists, so that we could fit more complicated models (beyond the independence model). If the two lists are truly independent, then the estimate of $\tilde{\beta}_{12}$ should be close to zero.

5 Estimating the number of police homicides in the United States

Using two data sources, a recent report by the Bureau of Justice Statistics (BJS) estimated that in the period 2003–2009 and 2011, there were 7,427 killings by a police officer, 2,103 of which were not reported to either of the two groups that created the lists used in their analysis (Banks et al., 2015). This estimate was obtained using multiple systems estimation applied to two lists: the Arrest-Related Deaths (ARD) database and the Supplementary Homicide Reports (SHR) released by the FBI. Because they only had access to two lists of killings, the authors implemented the standard two system estimator, which assumes list independence. However, this assumption is implausible because the dependence can be caused by either direct sharing of data or other common sources of information between the two lists, or by heterogeneity in probability of capture among members of the population. Thus it is of interest to test the sensitivity of the estimates to the presence of list dependence.

The BJS report's description of how the data were collected hints that the lists are plausibly positively correlated. For example, the BJS report states that 26 out of 50 states relied solely on public media to identify homicides for the ARD. Those homicides that were visible enough to have attracted media attention may be wealthier or otherwise more socially prominent relative to victims who did not attract media attention. People more likely to have been noticed and documented by the media may have been more likely to be documented by the SHR as well. Thus, a victim with high social visibility would be likely to appear on both lists. As a result, it is plausible that the estimate of 2,103 unreported killings underestimates the true number (we argued this in a blog post soon after the BJS report release). However, without additional information, it is difficult to provide a quantitative estimate of the magnitude of underestimation.

Experience with multi-list estimates in similar contexts in other countries also suggests that there is likely positive list dependence. HRDAG has used multiple systems estimation to estimate the number of homicides in many places, usually in the context of intra- or interstate conflict. In all of these examples, at least some of the killings are by police. The examples used here include homicides in Colombia,[1] Kosovo,[2] Guatemala,[3] Sierra Leone,[4] and Syria.[5] The Colombia data, in particular, offers a relevant comparison because it includes one list from the Colombian National Police and a second list coded from media and non-governmental reports of homicides, similar to the SHR and ARD databases used in the BJS report.

In contrast to the BJS report, each of HRDAG's estimates referenced here used more than two lists. From these projects, we have a large collection of estimates of pairwise list dependence to use in conducting a sensitivity analysis on the BJS estimate. Using data on homicides from these five countries, we estimate the pairwise list dependence in each of these multi-list models. We find that even across different conflict contexts and seemingly different data collection techniques, the distribution of pairwise list dependence is surprisingly stable.

We now apply these distributions of pairwise list dependence to the data on police homicides in the United States. The result yields estimates of the number of unrecorded police homicides under the assumption that the pairwise list dependence in the ARD-SHR projects is similar to list dependence in the projects in the countries considered. Because of the similarity of the distribution of pairwise list dependence across these different countries and time periods, the resulting distribution of police killings in the United States is remarkably unchanging when each of these six other countries is used as a surrogate. Our adjusted estimates indicate that, if list dependence between the ARD and SHR is similar to that in any of these other contexts, approximately 10,000 total police homicides occurred over the period of 2003–2009 and 2011 – leaving roughly half of the killings unobserved by both the ARD and the SHR.

5.1 Data on police homicides

We apply the method described in Section 4 to estimate the number of police homicides in the United States. The BJS report lists an aggregate estimate of 2,103 unreported police homicides. This number was arrived at by first partitioning, or stratifying, their incidents into subgroups such that

[1]See Lum et al. (2010). This data includes data from police, prosecutors, and media, among other sources, and is especially comparable to the ARD-SHR comparison in the BJS report.

[2]See Ball et al. (2002), online here.

[3]See Ball (1999), online in English here.

[4]The matched data are unpublished but derive from HRDAG's quantitative chapter in the report of the Truth and Reconciliation Commission, available here, as well as information from a subsequent report with additional data, see Guberek et al. (2006) online here.

[5]The estimates are unpublished, but the data are presented in a series of reports, the most recent of which is Price et al. (2014), online here.

the assumption that the incidents within each subgroup are homogeneous is tenable. In the BJS report, the data is stratified by several dimensions. Using the data within each stratum, they then perform multiple systems estimation, producing an estimate of the number of unrecorded homicides within each stratum. The estimate of the total number of unreported police homicides is then the sum of all of the individual stratum estimates.

Unfortunately, the full dataset used by the BJS analysts was not released. The only information that is available is the aggregate list overlap counts (as opposed to the stratum-by-stratum list overlap counts). The released data give $n_{11} = 1681$, $n_{10} = 1939$, and $n_{01} = 1704$. However, a cursory analysis reveals that the stratification did little to alter the final estimates. A two system log-linear independence model estimate on the aggregate counts gives an estimated number of unreported homicides of 1,966, which is only 137 short of the estimate obtained from aggregating stratum-wise estimates. Further, the confidence interval for the number of unreported police homicides is $[1831, 2131]$, which contains the BJS's stratified estimate. Given the close correspondence between the estimate using stratified data and the estimate using aggregate data, we proceed using only the aggregate list overlap count data.

5.2 Estimating pairwise list dependence using other, contextually similar data

To generate a plausible distribution of values of list dependence parameters, we estimate pairwise list dependence for each of several other homicide-related datasets. We stratify each country's data along a variety of dimensions (no disaggregation; disaggregation by region; by time; and by region and time). A few details about the stratification schemes used for each country's data are given in Table 1. These strata serve as the collection of populations in Section 4. We fit a Poisson regression model with all pairwise interaction terms (i.e. model 6) to each stratum's list intersection counts. Strata for which there are too many zeroes or not enough data (fewer than 60 records) are omitted from the analysis. From each fitted model, we compute all pairwise marginal odds ratios, $\tilde{\beta}_{12}$ as described in Section 4. These make up each country's distribution of list dependence parameters, shown in Figure 1.

We find that in all of the countries considered, the bulk of the mass of the distribution of estimates of the marginal log odds ratios falls to the right of zero, indicating positive list dependence. The proportion of list pairs exhibiting positive list dependence across all strata is given in Table 2.

Country	J	Temporal Stratification	Regional Stratification
Colombia	4	yearly	department
Guatemala	3	yearly	municipality
Sierra Leone	5	yearly	region
Kosovo	4	monthly	north/south/east/west
Syria	5	yearly	governorate

Table 1: Information about each country's data.

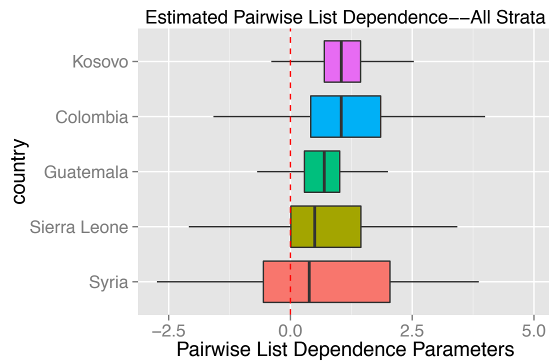

Figure 1: Distributions of estimated marginal log odds ratios between inclusion in lists of homicide data, as estimated from Kosovo, Colombia, Guatemala, Sierra Leone, and Syria. Estimates were obtained by first fitting a Poisson regression model with all pairwise interaction terms (i.e. model 6) to strata defined by Table 1, followed by equation 7 to obtain the marginal log odds.

5.3 Using contextually similar estimated pairwise list dependence to adjust estimates of police homicides in the US

Using the marginal log odds ratios estimated using each country's homicide data, we re-estimate the total number of police homicides in the US. As described in Section 4, we fit a two list model to the BJS data that includes both main effects and a pairwise interaction term, fixing the interaction term (β_{12}) to be our estimated values from our other populations. This produces a distribution of estimates under the assumption that list dependence in the US context is similar to that observed in each of the other countries. The results of this analysis are shown in Figure 2. The dashed line indicates the original estimates released by the BJS. Each box plot gives the distribution of estimates using the corresponding analysis' list dependence distribution. We conclude that if list dependence between the ARD and the SHR is similar to that seen in many other contextually similar analyses, the estimated number of police homicides in the US is likely in the range of 10,000. Our estimates presented here suggest that there are approximately 3,000 more

SY	SL	GT	CO	KO
0.68	0.76	0.87	0.85	0.93

Table 2: Proportion of lists exhibiting positive pairwise dependence by country.

undocumented police homicides than the BJS estimate. This increases the estimated proportion of unrecorded homicides from the original estimate of approximately 28% to 47%.

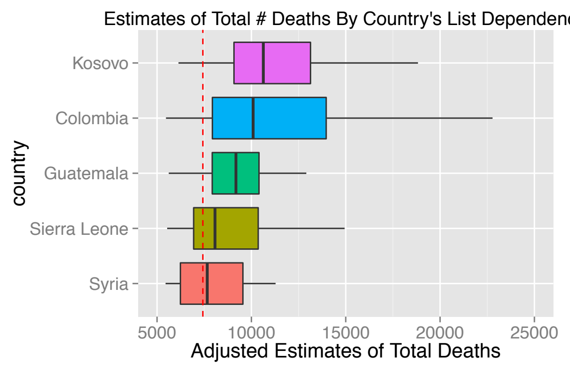

Figure 2: Distributions of estimates of police homicides during the period 2003–2009 and 2011, adjusted using marginal log odds ratios estimated from Kosovo, Colombia, Guatemala, Sierra Leone, and Syria. The dashed line indicates the original estimates from the Bureau of Justice Statistics report.

As mentioned in Banks et al. (2015), these list intersection counts only include jurisdictions that reported any data (about 70%). As such, these numbers should be interpreted as estimates of the number of people killed by police *in the reporting jurisdictions*. If the reporting jurisdictions are missing from the dataset not because there were truly no killings in those areas during this period, but instead because they chose not to report homicides by police, the true number of police homicides could be 30% higher than we have suggested here.

6 Conclusion

This analysis addresses the problem of unreported homicides. In our experience working in more than thirty countries experiencing violent conflict, some fraction of homicides are always hidden. HRDAG encourages the US authorities to record all homicides committed by police, and to make the lists of victims publicly available. Documenting deaths is a fundamental obligation of governments, and deaths resulting from government actions are perhaps the most important category of deaths to be recorded.

Even in the most rigorous documentation contexts, some homicides are likely be unreported. All lists of homicide victims should be tested for completeness using methods like those used in the BJS report. We welcome the BJS report as a critical step forward in the documentation of police homicides in the US.

References

Ball, P. (1999). Metodología Intermuestra. In *Guatemala: Memoria del Silencio, Vol. XII*. United Nations Commission for Historical Clarification, Guatemala.

Ball, P., Betts, W., Scheuren, F., Dudukovich, J., and Asher, J. (2002). *Killings and Refugee Flow in Kosovo, March–June, 1999*. American Association for the Advancement of Science and American Bar Association's Central and Eastern European Law Initiative, Washington, D.C.

Banks, D., Couzens, L., Blanton, C., and Cribb, D. (2015). Arrest-related deaths program assessment. Technical Report NCJ 248543, RTI International.

Bishop, Y. M. M., Fienberg, S., and Holland, P. H. (1975). *Discrete Multivariate Analysis: Theory and Practice*. MIT Press, Cambridge, MA.

Cormack, R. M. (1992). Interval estimation for mark-recapture studies of closed population. *Biometrics*, 48:567 – 576.

Cormack, R. M. and Jupp, P. E. (1991). Inference for Poisson and multinomial models for capture-recapture experiments. *Biometrika*, 78(4):911–916.

Guberek, T., Guzmán, D., Silva, R., Cibelli, K., Asher, J., Weikart, S., Ball, P., and Grossman, W. (2006). Truth and myth in Sierra Leone: An empirical analysis of the conflict, 1991–2000. Technical report, Human Rights Data Analysis Group.

Guzmán, D., Guberek, T., and Price, M. (2011). Unobserved Union Violence: Statistical Estimates of the Total Number of Trade Unionists Killed in Colombia, 1999-2008. Technical report, Benetech Human Rights Program.

International Working Group for Disease Monitoring and Forecasting (1995). Capture-recapture and multiple-record systems estimation I: History and theoretical development. *American Journal of Epidemiology*, 142(10):1047–1058.

Lang, J. B. (1996). On the comparison of multinomial and Poisson log-linear models. *Journal of the Royal Statistical Society, Series B: Methodological*, 58(1):253–266.

Lum, K., Price, M., Guberek, T., and Ball, P. (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).

Price, M., Gohdes, A., and Ball, P. (2014). Updated statistical analysis of documentation of killings in the Syrian Arab Republic. Technical report, Human Rights Data Analysis Group.

About HRDAG

The Human Rights Data Analysis Group is a non-profit, non-partisan organization[6] that applies scientific methods to the analysis of human rights violations around the world. This work began in 1991 when Patrick Ball began developing databases for human rights groups in El Salvador. HRDAG grew at the American Association for the Advancement of Science from 1994–2003, and at the Benetech Initiative from 2003–2013. In February 2013, HRDAG became an independent organization based in San Francisco, California; contact details and more information are available on HRDAG's website (`https://hrdag.org`) and Facebook page.

HRDAG is staffed by applied and mathematical statisticians, computer scientists, demographers, and social scientists. HRDAG supports the protections established in the Universal Declaration of Human Rights, the International Covenant on Civil and Political Rights, and other international human rights treaties and instruments. HRDAG scientists provide unbiased, scientific results to human rights advocates to clarify human rights violence. The human rights movement is sometimes described as "speaking truth to power:" HRDAG believes that statistics about violence need to be as true as possible, using the best possible data and scientific methods.

The authors thank Dr Shira Mitchell for her contribution to the statistical methods of this publication. The materials contained herein represent the opinions of the authors and editors and should not be construed to be the view of HRDAG, any of HRDAG's constituent projects, the HRDAG Board of Advisers, the donors to HRDAG or to this project.

[6]Formally, HRDAG is a fiscally sponsored project of Community Partners.