Assessing Claims of Declining Lethal Violence in Colombia*

Executive Summary: In this exploratory research note, we assess recent claims that violence in Colombia declined after the demobilization of paramilitaries. We show that these claims rest both on the overinterpretation of unadjusted data and on unsound causal inferences. We conclude that multiple data sources are needed to estimate the true rates of violence in Colombia after demobilization, and we suggest avenues for further research toward this end.

*This exploratory research note was prepared by Patrick Ball, Tamy Guberek, Daniel Guzmán, Amelia Hoover, and Meghan Lynch of the Human Rights Program of the Benetech Initiative.

Introduction

Two recent articles have received wide attention for their quantitative claim that lethal violence[1] in Colombia has declined following the paramilitary demobilizations.[2,3] Spagat and CERAC (undated) – though not González Peña and Restrepo (2006) – further make the unsupported claim that demobilization caused the decrease in lethal violence.[4] Neither the factual nor the causal claim can be justified by the evidence that the authors provide.[5]

The articles mentioned above use only one dataset to represent violence, the Colombian National Police registry of homicides (DIJIN). In this preliminary investigation, we find evidence consistent with the hypothesis that police reporting of lethal violence varies over time and across regions. The DIJIN database alone may be an unreliable measure of total homicides (in the technical sense, explained further below). We show further that if the potential underreporting biases were corrected, the rate of homicides after the demobilization might be greater than the rate before the demobilization. Simple counts of homicides in any database are more likely to reflect changes in data collection strategies and challenges than real changes in violence.

We also briefly address the assumption, implicit in Spagat and CERAC (undated) and González and Peña (2006), that homicides in the DIJIN database are "independent and identically distributed cases." If this technical assumption is false, then the regression models used to estimate the effect of demobilization are misspecified. Furthermore, if aspects of the paramilitary groups themselves determined when demobilization occurred, unmeasured attributes of the paramilitary groups could have an independent effect on the level of violence in Colombia. We conclude by pointing to directions for further research.

These results are preliminary. We are currently conducting the research outlined in the section on future research. We welcome comments, methodological suggestions, and new data sources.

[1] For the purposes of this paper, "lethal violence" refers to homicides where a body was found, because this is the dependent variable that Spagat and CERAC (undated) use. In the future, we hope also to analyze another form of lethal violence, namely disappearances. Disappearances are not included in the DIJIN dataset.

[2] Spagat, Michael and CERAC. undated. Colombia's Paramilitary DDR: Quiet and Tentative Success. Reprinted [in Spanish] in United Nations Development Programme, *Hechos de Callejón*.

[3] González Peña, Andrea, and Jorge Alberto Restrepo. 2006. Desmovilización de las AUC: Mayor Seguridad Humana? *UN Periódico* 97 (September).

[4] Spagat and CERAC state: "... the average demobilization... *lowers the homicide rate* by (a statistically significant) 13% in the area of operation of the demobilized group. A further calculation indicates that between 1400 and 2800 homicides *have been averted so far due to the cumulative effect of all the completed demobilizations.* Thus, paramilitary DDR has paid dividends for Colombia" (undated, 3-4; emphasis added).

[5] Despite requests, the authors have declined to provide the basic statistical results from their study (i.e., the regression parameters), so we are unable to comment further on their method.

Outline of Critiques

There are many problems with relying solely on one dataset. For example, the articles mentioned above gather their homicide data only from the Colombian National Police (the DIJIN dataset). In order to make their factual claim ("lethal violence has declined"), the authors assume that the Colombian police have records of all homicides that occur in Colombia. Furthermore, Spagat and CERAC's causal analysis ("lethal violence has declined as a result of paramilitary demobilization") assumes that the timing of demobilization was independent of the characteristics of paramilitary groups. Both of these assumptions are flawed. However, the following preliminary analysis considers only the measurement problem in depth, since accurate measurement necessarily precedes the testing of any causal claims. We find strong reason to doubt the factual claims advanced by both articles and the causal claim advanced by Spagat and CERAC (undated).

The analysis in this paper is exploratory. As a result, we neither reject nor present data consistent with the various claims about changing homicide patterns at this time. Rather, this first round of analyses merely demonstrates that available data are insufficient to support the arguments of the articles cited above. We also present a research agenda capable of rigorously testing such claims.

In our experience investigating mass killings in Guatemala,[6] Kosovo,[7,8,9] Perú,[10] and Timor-Leste,[11] we have found that only a fraction of homicides are reported to governmental or non-governmental organizations in conflict situations. There are many reasons why homicides may not be reported or registered by authorities at all or may be recorded at variable rates:

[6]Ball, Patrick. 1999. Metodología intermuestra. Guatemala: Memoria del Silencio. Vol. 12. CEH. Reproduced in English in Patrick Ball, Herbert Spirer, and Louise Spirer, eds. 2000. *Making the Case: Investigating Large Scale Human Rights Violations Using Information Systems and Data Analysis*. Washington, DC: AAAS.

[7]Ball, Patrick (with the American Bar Association-Central and East European Law Initiative). 2000. Political Killings in Kosova/Kosovo, March-June 1999. Washington, DC: ABA/CEELI-AAAS.

[8]Ball, Patrick, Wendy Betts, Fritz Scheuren, Jana Dudukovic, and Jana Asher. Washington, DC: AAAS and ABA/CEELI. 2002. Report presented as expert testimony to International Criminal Tribunal for the Former Yugoslavia.

[9]Ball, Patrick, Meghan Lynch, and Amelia Hoover. 2007. "Revisiting 'Killings and Migration in Kosovo:' Responses to additional data and analysis." Report to be presented as expert testimony to the International Criminal Tribunal for the Former Yugoslavia.

[10]Ball, Patrick, Jana Asher, David Sulmont, and Daniel Manrique. 2003. 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. Report to the Peruvian Commission for Truth and Justice (CVR). Also published as Anexo 2 (Anexo Estadístico) of CVR Report, 28 August 2003. Washington, DC: AAAS.

[11]Silva, Romesh, and Patrick Ball. 2006. The Profile of Human Rights Violations in Timor-Leste, 1974-1999. Report by the Benetech Human Rights Data Analysis Group to the Commission on Reception, Truth and Reconciliation (CAVR). Available online at: `http://www.hrdag.org/resources/publications/Benetech-CAVR-statistical-report.pdf`. Accessed 13 Feb. 2007.

- perpetrators may hide the bodies from police to avoid prosecution;
- severe crimes often occur in areas so remote or dangerous that police access is difficult;
- police budgets and staff may change, making the police more or less able to register crimes;
- the police may have different operating procedures in different locations, or it may change its operating procedures over time;
- the police may have been intimidated by the perpetrators;
- the victims' families may hide the crime in order not to draw attention to themselves;
- members of the police may themselves have been the perpetrators or may be collaborating with the perpetrators;
- perpetrators' tactics may have changed, while police reporting methods remain the same.[12]

Some or all of these factors are likely to be present in various police jurisdictions in Colombia. We know this because decades of criminological research in North America have shown that several factors affect how much police know and record about total crime being committed.[13,14] Thus, analyses and conclusions should not be based solely upon police reports treated as absolute fact. This is especially important in light of the doubts that the Colombian police itself has expressed about the reliability of its own statistics.[15] Indeed, *any* single dataset, regardless of its source, is almost certainly incomplete.[16] We can only know how and in which ways a dataset is incomplete by comparing it to other (also incomplete) datasets. It is irresponsible to draw conclusions from any single dataset, whether it be the DIJIN data or any other.

[12]For example, disappearances may increase as homicides (for which a body is found) decrease. If police reporting procedures remain the same (defining a homicide as a case in which the body is found), then this substitution effect may mean that even if the same number of people is killed or disappeared, it may be reported only as a decrease in the total number of homicides.

[13]See, e.g., Kennedy, Leslie W. and David Veitch. 1997. Why Are Crime Rates Going Down? A Case Study in Edmonton. *Canadian Journal of Criminology* 39. This analysis finds that the most important factor influencing reported crime rates are administrative decisions taken to alter the access to police resources for the public.

[14]Note also the United States Federal Bureau of Investigation (FBI) admission that "crimes reported to police" are a poor proxy for total crime: http://www.fbi.gov/ucr/ucrquest.htm, "How can I compare the UCR Program's findings with that of the NCVS?" Accessed 7 Feb. 2007.

[15]See "Las extrañas cifras de seguridad de la Alcaldía." 2 Sep. 2006. *El Tiempo.* and "DIJIN explica las razones de las inconsistencias en cifras de seguridad de Bogotá." 8 Sep. 2006. *El Tiempo.*

[16]Here we are referring to datasets that compile information known to the collector, formally called convenience samples. Probability samples or formal experimental data would not be subject to the same limitations.

Examining the Reliability of DIJIN Statistics

DIJIN (Dirección Central de Policía Judicial, or the Central Directorate of Criminal Police) is a directorate within Colombia's Policia Nacional. DIJIN is responsible for judicial and criminal investigations to support the administration of justice; as such, it is the primary repository and disseminator of official homicide records. Thus, the homicides in the DIJIN database are those known to, and reported by, regional police command posts. While it is an important source of information, the DIJIN database may be, for reasons stated above, unlikely to reflect true patterns of lethal violence in Colombia.

Reliability

In order to draw conclusions about changes in the true homicide rate, Spagat and CERAC (undated) and González Peña and Restrepo (2006) rely on the relationships between homicide rates across time and space, as reflected in DIJIN statistics. However, to draw plausible conclusions from the data that they use requires the strong assumption that the Colombian Police always register a constant percentage of all homicides. While this is theoretically possible, it is far more likely that the rate of registration varies (see above and footnotes [6], [7], [8], [9], [10], [11]): in some periods and places, the police register almost all homicides, while in other periods and places, they register relatively few. This variation directly affects how many homicides are reported. If it became more difficult to record homicides for any reason following paramilitary demobilization, then the data would be systematically biased downward, toward a spurious finding of decreased lethal violence.

In the theory of measurement, "reliability" refers to the ability to obtain the same (or very similar) results from repeated measurements of the same object, whether that object be the weight of an atom, the standardized test performance of a student, or the number of murders in Colombia during a specific period of time. In order to claim that lethal violence is declining based on a single dataset, researchers would need to demonstrate that the dataset were reliable in this formal sense. If reporting rates declined while violence remained high, then higher rates of violence in the real world could appear stable or declining in a single dataset. Such a measurement of homicide would be unreliable, because another sample (an additional, independent dataset) would be likely to yield different results.

In order to assess the reliability of DIJIN data, we would examine cases of homicide reported in a hypothetical complete data source and determine the extent to which cases reported in the DIJIN database overlapped with those in the complete data. Low match rates[17] would indicate underreporting, while changing match rates would imply low reliability (in the formal sense). However,

[17] See below in Quantitative Analysis. "Match rate" refers to the extent to which cases reported in two data sources overlap. More specifically, it is reported below (see Tables 1 and 2) as the percentage of homicides sampled from an independent data source that are also reported in the DIJIN database.

no such perfectly complete dataset is available: *all datasets are incomplete*. Instead, we base our preliminary findings on comparisons between DIJIN data and an independent source (data from the Colombian Commission of Jurists, CCJ).

Quantitative Analysis

By using more than one dataset, it is possible to estimate the level of underreporting in a single dataset.[18] Once the level of underreporting – in this case, the estimated gap between the total number of homicides and the number of homicides reported to the police – has been determined, we can employ statistical techniques to correct for this underreporting. Because the present analysis relies on only two sources of data (DIJIN and the CCJ), the ultimate source of unreliability cannot be determined at this stage. However, we can state definitively that the DIJIN data (like any other single collection of data) do not provide a comprehensive picture of the situation.

In this first analysis, we are particularly concerned with the conclusions about homicide trends drawn by Spagat and CERAC (undated) and González Peña and Restrepo (2006) from uncorrected DIJIN data. We selected two test areas (Antioquia and Nariño) to match the cases presented in Spagat and CERAC (undated). Our third area, Cauca, was of special interest because it was among the first departments to experience paramilitary demobilization.[19] Each of these three areas underwent demobilizations at different times. For each area, we randomly selected homicides appearing in the CCJ database in the periods before and after the first paramilitary demobilization. The start and end dates for each period in each area are listed below:

- Antioquia Pre-demobilization period: 1 Jan 2003 - 25 Nov 2003
- Antioquia Post-demobilization period: 26 Nov 2003 - 31 Aug 2006
- Cauca Pre-demobilization period: 1 Jan 2003 - 7 Dec 2003
- Cauca Post-demobilization period: 8 Dec 2003 - 31 Aug 2006
- Nariño Pre-demobilization period: 1 Jan 2003 - 30 July 2005
- Nariño Post-demobilization period:[20] 31 July 2005 - 31 Aug 2006

[18] ChandraSekar, C. and W.E. Deming. 1947. On a method of estimating birth and death rates and the extent of registration. Journal of the American Statistical Association 44, 101-115. See also Marks, Eli S., William Seltzer and Karol Krótki. 1974. *Population Growth Estimation: A Handbook of Vital Statistics Measurement.* New York: The Population Council.

[19] Note, however, that neither the paramilitary organization nor the demobilization effort in Cauca is necessarily representative of those in other departments, underscoring the need for further analysis covering all departments.

[20] In Nariño, we used only 50 homicides in the post-demobilization period, because those homicides represented the entirety of the CCJ database for that region during the specified time period.

We then compared the homicide records sampled from the CCJ data to the police data using "loose matching" criteria.[21]

The results reported here are based on overlap or "match" rates, which are determined by deaths that appear in both datasets.[22] Higher match rates mean that a greater percentage of the CCJ data was recorded by the police and hence included in the DIJIN homicide list. Conversely, lower match rates mean that a lower percentage of the CCJ data was registered by the police and a lower percentage appears on the DIJIN list. If match rates change over time, we conclude that at least one data source displays low reliability. Our method is described in greater technical detail in the Data and Methods Appendix below.

Number of Matches, Pre- and Post-Demobilization, by Department

The tables that follow summarize the results of the matching process described above (Table 1), as well as estimates of the statistical significance of those results (Table 2).

Table 1: Matching between CCJ and DIJIN Datasets, by Department and Demobilization Period

Statistic	Antioquia		Cauca		Nariño	
	pre	post	pre	post	pre	post
CCJ database observations	379	503	96	199	127	50
CCJ sample size	75	75	75	100	75	50
CCJ matched to DIJIN data	49	38	45	49	54	25
CCJ match rate	0.65	0.51	0.60	0.49	0.72	0.50
Adjusted standard error	0.05	0.05	0.03	0.04	0.03	0

In Antioquia, the CCJ database contained 379 total homicides for the pre-demobilization period and 503 total homicides in the post-demobilization period. We sampled 75 cases from each period for matching. Forty-nine of the 75 homicides in the pre-demobilization CCJ sample were also found in the DIJIN dataset, a raw match rate of 65% (with margin of error, 55%-75%).[23] In the

[21] We compared the records by first name, middle name, first last name, second last name, date of death, department, municipality, and sex. By applying the same match critiera across pre- and post-demobilization periods, we created a fair test of the hypotheses.

[22] Reasoning about match rates, as well as point and variance estimators can be found in Bishop, Yvonne M. M., Stephen E. Fienberg, and Paul H. Holland. 1975. *Discrete Multivariate Analysis: Theory and Practice.* Cambridge, MA: MIT Press, especially Chapter 6. For a discussion of the use of this technique in the adjustment of censuses, see Robinson J. G., Ahmed B., das Gupta P. and Woodrow K. 1992. Estimation of Population Coverage in the 1990 United States Census Based on Demographic Analysis. *Journal of the American Statistical Association*, 88(423), p. 1061-1071.

[23] The margin of error (MoE) is easily calculated from Table 1. $MoE = \pm(1.96 \times ase)$, where ase is the adjusted standard error. Adjusted standard error is calculated as $ase = se \times fpc$, where se is the standard error and fpc is the finite population correction. Note that the standard error is small, because the sample size (50-100 homicides for each department) is large relative to the population (total homicides in the CCJ database). Standard error is calculated as $se = \sqrt{m(1-m)/(n-1)}$, where m is the raw match rate and n is the number

post-demobilization period, only 38 of 75 homicides in the CCJ sample were matched to DIJIN records, a raw match rate of 51% (with margin of error, 41%-61%). Even considering the margin of error, the difference between the pre-demobilization and post-demobilization reporting rates is statistically significant at the 95% level (see Table 2).

Results are similar for our other two replication cases, Cauca and Nariño. In Cauca, the raw match rate for the pre-demobilization period was 60% (54%-66%), which declined to 49% (41%-57%) after demobilization.[24] In Nariño, we observed an even more striking decline in match rates, from a pre-demobilization raw match rate of 72% (66%-78%) to a post-demobilization raw match rate of only 50%.[25] As in Antioquia, these pre-post differences are highly statistically significant. This finding is summarized in Table 2.

Table 2: Significance tests of pre- and post-DDR matching rates, by Department

Statistic	Antioquia	Cauca	Nariño
pre-post match level difference	0.15	0.11	0.22
standard error of difference	0.07	0.04	0.03
p-value of the difference	0.02	0.01	0

Returning to the case of Antioquia to interpret Table 2, we find that the difference between the pre-demobilization and the post-demobilization match rate (15%) is statistically significant. The "p-value of the difference" refers to the probability that we would measure a 15% difference in a similar sample if the true difference were zero. This probability is only 0.02 (2%), indicating that if we repeated the sampling 100 times, in 95% of the samples we would expect to identical substantive conclusions. Even stronger results hold for Cauca and Nariño, where the relevant probabilities are 1% and 0% (to two significant digits), respectively – highly significant results.

The results reported in Tables 1 and 2 indicate that, regardless of the true number of homicides following demobilizations in Antioquia, Cauca and Nariño – which is unknown and can only be estimated – the police reported a significantly smaller percentage of homicides in the CCJ sample during the post-demobilization period. If the CCJ reporting rate were relatively stable between periods, then we would further conclude that the police also recorded a significantly smaller proportion of *total* homicides in the post-demobilization period.

of CCJ cases sampled during the period. Finite population correction is calculated as $fpc = \sqrt{(N-n)/(N-1)}$, where N is the total number of cases in the CCJ database during the period and n is the number of CCJ cases sampled during the period. All calculations are reported to two significant digits.

[24] Note that in order to calculate estimated match rates with greater precision, 100 cases were sampled from the CCJ dataset for Cauca in the post-demobilization period.

[25] Note that in the case of post-demobilization Nariño, 50 cases represents the entirety of the CCJ dataset for the period (see fn 15 above); thus, $MoE = se = 0$.

Determining Causality

As noted above, Spagat and CERAC (undated) claim that "[t]he DDR process takes on much of the character of a controlled experiment due to great variation in the times and places of demobilizations... The average demobilization, of which there have been 37 so far, lowers the homicide rate by (a statistically significant) 13% in the area of operation of the demobilized group."[26] As we have shown above, there is little reason to be confident that decreases in the reported homicide rate reflect decreases in the actual homicide rate. In this section, however, we take the ostensibly lower homicide rate as given and ask: if the true homicide rate had, in fact, decreased, would the evidence be strong enough to support the claim that the decrease was *caused by* the demobilization process?

This hypothetical question is important for analyses of the Colombian conflict. If demobilization did not lower rates of homicide, then we would be forced to reconsider the effectiveness of demobilization as a policy.

In Spagat and CERAC's (undated) causal assertion, the crucial assumption is that the DDR process resembles a controlled experiment, in which departments are randomly assigned to the "treatment" of paramilitary demobilization. However, in order to claim that "the average demobilization... lowers the homicide rate," analysts must show that they can separate demobilization's average effect on lethal violence from other factors' effects on lethal violence. If the initiation of DDR processes were not assigned randomly (as subjects are assigned to treatment and control groups in an experiment) then these effects cannot be separated. As a result, any argument about demobilization's causal effects (on homicide rates or anything else) would be seriously weakened.

In order to separate the average effect of demobilization, an analyst must show that variation in time and place of demobilization affects homicide rates *only through the demobilization process.*[27] Unfortunately, time and place of demobilization are associated with other factors that may also affect homicide rates, including paramilitary group characteristics, the conflict situation, or the presence of Colombian state forces in the area. These potentially confounding factors, among others, are missing from Spagat and CERAC's controls.

Many scenarios are plausible, but consider the following: if demobilization is more likely in places where police and other state forces are already present and effective, then the apparent success of demobilization identifies areas with

[26] Spagat and CERAC (undated), pp.3-4.

[27] The statistical counterpart to a natural experiment research design is known as an instrumental variable. If we are interested in the effect of X (and only X) on Y, but we cannot isolate the effect of X (assume, for example, that M and N also affect Y), then one way to find X's independent effect is to think of another variable, Z, that directly affects X *without* otherwise affecting Y. In this case, Z is the instrumental variable. For a more technical discussion see Greene, William. 2002. *Econometric Analysis (5th ed.),* chapter 9. Upper Saddle River, NJ: Prentice Hall. For an application of instrumental variables reasoning to conflict dynamics see, e.g., Miguel, Edward, Shanker Satyanath, and Ernest Sergenti. 2004. Economic Shocks and Civil Conflict: An Instrumental Variables Approach. *Journal of Political Economy* 112, 725-753.

effective police forces, rather than the success of the demobilization policy *per se*. There exist many such possible scenarios, and therefore the assumption of quasi-randomness cited above should be strongly questioned. Researchers must explicitly account for correlations between armed group characteristics, conflict dynamics, and demobilization processes in any study that claims to emulate experimental methods in the Colombian conflict.[28]

Discussion and Future Research

The agenda laid out in this exploratory research note generates many paths for future research. We have shown that a given dataset's coverage of total violence must always be considered partial. Using multiple, independent datasets can provide a basis for estimating the true patterns of violence.[29] Even in a country like the United States, which has not for many decades experienced civil conflict and has substantial resources to devote to data collection, "crimes reported to the police" differs from estimates of the total amount of crime. The Federal Bureau of Investigation notes, for example, that "changes in police procedures, shifting attitudes towards crime and police, and other societal changes can affect the extent to which people report and law enforcement agencies record crime."[30] Indeed, in an earlier article, Spagat and his co-authors note the problems with the DIJIN data: "In the case of Colombia, existing data sets have the above [problems relating to all datasets of situations of conflict] and other problems. ... The main basis for data inclusion [in the DIJIN dataset] is reports to police authorities or events known to police authorities in townships, *excluding events that are not reported, particularly those that occur outside of institutional presence*. (emphasis added)"[31] Given the theoretical flaws affecting analysis based on any single dataset, and given the empirical observation that the DIJIN data seem to vary substantially in their coverage, we consider it irresponsible to draw any conclusions from the DIJIN dataset alone.

A first step in future research should be to use multiple datasets to estimate total violence over time and space. To do this correctly and responsibly requires a great deal of care. For example, the scale of the Colombian case means that duplicate cases cannot be identified by hand-matching. Multi-year datatsets contain tens or hundreds of thousands of recorded homicides. For consistency across all these data, the matching process must be computer-generated using

[28] It is important to note that the CERAC authors' models included several demographic and other control variables. However, they did not include specific measures of the aspects of the paramilitary groups themselves and their relationships with other armed actors that are at issue in this argument.

[29] Ball, Patrick. "Making the Case: The Role of Statistics in Human Rights Reporting." *Statistical Journal of the United Nations.* ECE 18. 2001. 163-73. See also Patrick Ball, *Who Did What to Whom?*" 2nd ed. HRDAG working paper, 2007. Forthcoming from Benetech.

[30] Appendix IV, "The Nation's Two Crime Measures." *Crime in the United States.* 2003. <http://www.fbi.gov/ucr/03cius.htm>. Accessed Feb. 7, 2007.

[31] See Restrepo, Jorge, Michael Spagat, and Juan F. Vargas. "The Dynamics of the Colombian Civil Conflict: A New Data Set." Discussion Paper Series. Royal Holloway University of London. Copyright 2004.

modern deduplication techniques.[32] The Human Rights Data Analysis Group at the Benetech Initiative is currently working on this project.

The second step should be to extend the multiple-dataset analysis to the whole country. Estimating levels of underreporting using multiple data sources will allow a far more accurate assessment of important variables in Colombia's conflict. Choosing a small number of departments for analysis risks selection bias – the analyst could choose departments that fit his or her explanation and exclude other departments that did not fit. The extension of the present analyses to the entire country is an important first step in the evaluation of purported causal relationships between conflict dynamics (including DDR processes) and levels of violence.

More generally, causal analysis can only follow the statistically sound measurement of changes in levels of lethal violence in Colombia. Furthermore, causal analysis must not purport to draw conclusions that the data do not support.[33] Sound causal analysis can only grow out of sound measurement practices, but even after careful measurement, causal analysis must operate within the limits of the data.

Conclusion

Existing analyses that claim that violence is declining in Colombia suffer from logical and technical problems. This paper has examined the DIJIN data collected by the Colombian police and has found that this dataset may not reflect the true rate of homicides. Rather, the fraction of all homicides reported by the CCJ that are also reported by police varies over time and place. In particular, as we have shown above, the rate at which the police report homicides in the pre- and post-demobilization periods may differ in a statistically and substantively significant way. Thus, claims that homicides have decreased in the post-demobilization period likely rest on analytic artifacts that resulted from changes in the data collection process. It may be that police reporting has declined, while the true homicide rate has remained the same or has even increased. Our findings show the necessity of verifying that data are reliable – that they reflect true patterns of violence, rather than spurious patterns generated by mechanisms such as changes in reporting.

[32] An introduction to machine learning techniques for database deduplication and matching is Bilenko, Mikhail, William W. Cohen, Stephen Fienberg, Raymond J. Mooney, and Pradeep Ravikumar. "Adaptive Name-Matching in Information Integration." *IEEE Intelligent Systems*, 18(5), pp. 16-23, September/October 2003. See also Witten, Ian H. and Eibe Frank. 2005. *Data Mining: Practical Machine Learning Tools and Techniques*, 2d ed. San Francisco: Elsevier, Inc.

[33] Even if Spagat and CERAC's (undated) data were reliable, their graphs show a decrease in violence in areas with no paramilitary activity occurring simultaneously with the reported decrease in areas with paramilitary activity. Alternate hypotheses must be considered – in Spagat and CERAC's (undated) case, this would have involved considering an unknown factor which might have accounted for a decrease in reported violence both in paramilitary and non-paramilitary zones.

Appendix: Data and Methods

Data from DIJIN

DIJIN (Dirección Central de Policía Judicial, or the Central Directorate of Criminal Police) is a directorate within Colombia's Policía Nacional. DIJIN is responsible for judicial and criminal investigations to support the administration of justice; as such, it is the primary repository and disseminator of official homicide records. The National Police began keeping a database with names of homicide victims in 2003. They collect their data nationwide through their agents and posts throughout the country. However, importantly, the only homicides reported in the DIJIN database are those that are known to, and reported by, regional police command posts. The National Police data include homicides due to many causes, including death in combat. For this study, we used all the National Police homicide data for three deparatments between January 1, 2003 to August 31, 2006.

Data from the Colombian Commission of Jurists

The research area of the Colombian Commission of Jurists (CCJ) has created a database of sociopolitical violence with the purpose of keeping a daily and permanent record on human rights violations and breaches of humanitarian law that take place in Colombia. By "sociopolitical violence," the CCJ refers to acts that constitute attempts against life, personal integrity and personal freedom caused through the abuse of power by agents of the State; those that are politically motivated; those that derive from discrimination against socially marginalized persons; or those caused by the internal armed conflict.

At present, the information processed by the CCJ stems from the following sources: twenty national and regional newspapers; two weekly news magazines with nation-wide circulation; direct complaints gathered by the CCJ; complaints made to other human rights organizations, national as well as regional, that monitor the situation of human rights and humanitarian law; information submitted by the national authorities (Public Defender's Office, Attorney General's Office, Prosecutor General's Office, State Forces) and the publication Noche y Niebla (Night and Fog) of the CINEP/Justicia y Paz Data Bank.

Method

Samples were drawn randomly from the CCJ list of political killings in six strata: in Antioquia, Cauca, and Nariño, both before and after the demobilization. The samples were matched to the DIJIN database. In all strata, the sample size constituted a substantial fraction of all homicides in the CCJ database, so the sampling error includes the finite population correction.

The matching criteria were as follows: A pair of records were considered the same if they had the same names and a slightly different date or municipality. If they had the same place and date, one of the four names could be slightly

different (often due to spelling errors in the database). Most dates that did not match failed to do so because one digit in the date format DD/MM/YYYY was different (for example, day and month were inverted, which suggests data coding errors). Missing data were treated differently than contradicting data. For example, if a record in one dataset had no middle name registered, this was considered less severe than if the middle names are available and were different.

About the Authors

Patrick Ball, Ph.D., is the Director of the Human Rights Program and Chief Technical Officer at the Benetech Initiative. Since 1991, Dr. Ball has designed information management systems and conducted statistical analysis for large-scale human rights data projects used by truth commissions, non-governmental organizations, tribunals and United Nations missions in El Salvador, Ethiopia, Guatemala, Haiti, South Africa, Kosovo, Sierra Leone, Perú, Timor-Leste, Sierra Leone, and Chad. Dr. Ball is currently involved in Benetech projects in Sri Lanka, Colombia, Burma, Liberia, Guatemala and in other countries around the world.

Tamy Guberek, B.A., is the Latin America Field Coordinator for the Benetech Human Rights Program, managing projects in Colombia and Guatemala. Ms. Guberek supports partners with human rights information management systems and data analysis. She has also contributed to the descriptive statistical analysis for Benetech analysis of violence in Sierra Leone. She received her B.A. in International Relations and Peace and Justice Studies from Tufts University.

Daniel Guzmán, B.S., is a statistical consultant for the Benetech Human Rights Program. He has contributed to project design and data analysis for Colombia, Guatemala and Sierra Leone. Mr. Guzmán served as the teaching assistant for Benetech's 2005 course on "Measuring Human Rights Violations" in Bogotá, Colombia. He received his B.S. in Statistics from the National University of Colombia.

Amelia Hoover, M.A., currently of the Human Rights Program at the Benetech Initiative, is a Ph.D. student in political science at Yale University. Her dissertation research focuses on human rights abuses during armed conflict, including covariation between lethal and non-lethal forms of violence and the effects of armed group command and control structures on patterns of violence.

Meghan Lynch, M.A., currently of the Human Rights Program at the Benetech Initiative, is a Ph.D. student in political science at Yale University. She studies political violence, particularly mass violence against civilians, with a regional focus on Africa. Her contribution to this work has been supported under a National Science Foundation Graduate Research Fellowship.

Patrick Ball designed the analysis, reviewed the calculations, and drafted the report. Tamy Guberek and Daniel Guzmán made contact with the National Police and the CCJ to get data. Ms. Guberek did the hand-matching and contributed to the writing. Mr. Guzmán drew the stratifed sample and calculated the rates before and after demobilization. Amelia Hoover and Meghan Lynch

contributed to the analysis, writing, and editing.

Our thanks to Beatriz Vejarano for research and translation support, to the Colombian National Police and the CCJ for data access, and to Ana María Gómez López, Michael Reed Hurtado, Fritz Scheuren and Romesh Silva for helpful comments.

About the Benetech Human Rights Program

HRDAG (the Human Rights Data Analysis Group) designs and builds information management solutions and conducts statistical analysis on behalf of human rights projects. With our partners, we make scientifically-defensible arguments based in rigorous evidence (`http://www.benetech.org`, `http://www.hrdag.org`).

This project was funded by core support to the Benetech Initiative from the Omidyar Network, the Skoll Foundation, the John D. and Catherine T. MacArthur Foundation, and The Oak Foundation.

The materials contained herein represent the opinions of the authors and editors and should not be construed to be the view of the Benetech Initiative, any of Benetech's constituent projects, the Benetech Board of Directors or the donors to Benetech.

480 S. California Ave., Suite 201
Palo Alto, CA 94306-1609
tel: +1 650-475-5440
fax: +1 650-475-1066
Email: info@benetech.org
Web: http://www.benetech.org