Comments to the article "Is Violence Against Union Members in Colombia Systematic and Targeted?"

Megan Price*and Daniel Guzmán†

Benetech Human Rights Program

Benetech
Technology Serving Humanity

May 28, 2010

1 Executive Summary

For decades, thousands of union leaders and members have been killed, disappeared and threatened in Colombia. Yet magnitude, patterns, and causes of anti-union violence are debated. Over the past two years, that debate and the attention paid to it has intensified, particularly as countries negotiating free trade agreements with Colombia, such the United States, have made union violence an explicit obstacle to finalizing agreements.

In November 2009, two Colombian academics, Daniel Mejía and María José Uribe, from the *Centro de Estudios sobre Desarollo Económico* (CEDE) at the *Universidad de los Andes* published a study entitled, "Is Violence Against Union Members in Colombia Systematic and Targeted?" This paper concludes that "...on average, violence against unionists in Colombia is neither systematic nor targeted." (p.1)

Given the political, economic and social importance of this debate, any study that makes claims about patterns and magnitude of union violence in Colombia requires the highest level of precision and scientific rigor. Therefore, in our response, we present – in technical and methodological detail – the reasons we find the conclusions in Mejía and Uribe's study to be overstated. In short, we believe that weaknesses in the data, in the model choice, and in the model interpretation used in

*Megan Price, Ph.D., is a statistician with the Benetech Human Rights Program. She has contributed statistical analyses to projects in Guatemala and Colombia. Dr. Price earned her Ph.D. in Biostatistics from the Rollins School of Public Health at Emory 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 projects in Colombia, Guatemala and Sierra Leone. He received his B.S. in Statistics from the National University of Colombia.

Mejía and Uribe's study, all raise serious questions about their strong causal conclusions.

Based on our careful review and critique, we conclude that Mejiía and Uribe's study does not resolve the question, "is violence against union members in Colombia systematic and targeted?" for following reasons:

- Unknown under-registration. The Mejía and Uribe study uses convenience sample data as the basis for its claims. These data are based on available, observable reports on union homicides and union activity which have been collected without a scientific random selection method. These data cannot be relied upon to represent an underlying larger population or to accurately describe patterns over time and space.
- Possible violation of model assumptions. The statistical methods used in the Mejía and Uribe study to evaluate the relationship between union homicides and "union activity" are based on commonly used least squares regression and instrumental variables analyses. These methods rely on very strong assumptions. However, we do not believe that the study adequately addresses the data's potential violations of these assumptions and the potential ramifications on the estimates when these assumptions are violated. Violations of these assumptions could change the magnitude of parameter estimates (used to quantify the relationship between union activity and union violence) and the significance of parameter estimates (used to determine the presence or absence of any relationship).
- Uncorrelated errors. One of the main modeling assumptions which we believe these data violate, uncorrelated errors, has a direct affect on significance tests used in the Mejía and Uribe study. Therefore we are highly skeptical about the reliability of the conclusion that union activity and union violence are not significantly associated.
- Poor quality models. The descriptive and analytic results presented in Mejía and Uribe's study indicate that union member homicide rate is a highly variable outcome measure. This variation results in poor-quality models. We find that control variables in the various model formulations show inconsistent and indeed reversed effects, suggesting problems with the model specification, with the data on homicides, or both.

The poor quality of these models, the unknown under-registration inherent in the data, and the questionable modeling decisions, mean that the strong conclusions in Mejía and Uribe's study are unsupported by the analyses. We point out that unchecked, those conclusions distort the truth about violence against unions and can mislead important social, economic and political decisions in Colombia.

In addition to questioning Mejía and Uribe's conclusions about whether violence against unions is systematic and targeted, we believe that the broader question about overall patterns and magnitude of union homicides in Colombia is still unanswered. We plan to continue this scientific and statistical debate which is relevant for current trade negotiations, and more importantly, for clarification of the historical truth about the victims of human rights violations in Colombia.

2 Introduction - The Importance of this Methodological Debate

For decades, thousands of union leaders and members have been killed, disappeared and threatened in Colombia. Union affiliates have suffered a great deal of violence, yet magnitude, patterns, and causes of anti-union violence are debated. Over the past two years, that debate and the attention paid to it has intensified, particularly as countries negotiating Free Trade Agreements with Colombia, such the United States, have made union violence an explicit obstacle to finalizing agreements. Several projects are currently working to analyze anti-union violence in Colombia.

For years, the most active organizations reporting about union violence have been the Medellín-based National Union School (*Escuela Nacional Sindical* (ENS)), the CUT (*Central Unitaria de Trabajadores*), the Vice President's Human Rights Observatory and the Colombian Commission of Jurists (CCJ). These reports are based on observed data on acts of violence and legal proceedings. More recently, new actors have come on the scene with new studies and new methodologies using the existing data. One such study was published by Daniel Mejía and María José Uribe, from the *Centro de Estudios sobre Desarollo Económico* (CEDE) at the *Universidad de los Andes*, in November 2009. Their paper, "Is Violence Against Union Members in Colombia Systematic and Targeted?" makes strong quantitative claims about the patterns and magnitude of union violence in Colombia.

Given the political, economic and social importance of this debate, we believe that any study that makes claims about patterns and magnitude of union violence in Colombia requires the highest level of precision and scientific rigor. In our role as the statistical advisors to the CCJ, we coordinated an initial response to Mejía and Uribe [Nov. 2009]. The content of this response was taken up by the ENS in their own response to the Mejía and Uribe study. The authors responded to those comments [Mejía and Uribe, Dec. 2009], and we now welcome the opportunity to continue this very important methodological discussion. In this paper, we present our concerns about the causal conclusions presented in Mejía and Uribe [Nov. 2009], taking into consideration the available data, model choices, and model interpretations. We are especially gratified that this is a methodological rather than ideological debate about union violence. In the following sections we will make specific references to tables, figures, and page numbers in Mejía and Uribe [Nov. 2009].

3 Background

The paper by Mejía and Uribe concludes that "...on average, violence against unionists in Colombia is neither systematic nor targeted" (p.1). This conclusion is based on both descriptive and empirical analyses. We outline our concerns in full in the following sections, but briefly, we see two main problems with the analysis and subsequent conclusions presented in this paper. First, given the type of data available (convenience sample) and structure of the data (panel, or repeated measures), we do not agree with the methodological choices in Mejía and Uribe [Nov. 2009]. Second, given the choices made in Mejía and Uribe [Nov. 2009], we interpret the results differently and draw more skeptical conclusions. We believe that the paper inadequately addresses the assumptions required by the methods, that it presents overstated findings, and that the analyses do not support strong causal claims.

The Mejía and Uribe paper begins by describing the data used on the number of union member homicides in Colombia between 1986 and 2008 and compares patterns in these data over time to patterns of total homicides and the homicide rate among vulnerable groups in Colombia. At this point, it is important to keep in mind that these are data from convenience samples[1] and as such, should not be assumed to represent the entire population of union members or union member homicides. Using convenience samples that capture an unknown fraction of the population to draw conclusions about the entire population, or patterns over time and space, can lead to biased conclusions. This concern is elaborated in Section 4.

Based on these comparisons, Mejía and Uribe [Nov. 2009] describes "... a continuous decrease in violence against union members and union leaders in Colombia." (p.7) However, as outlined in more detail in section 4, we disagree with this description of the data. We believe the data on union homicides presented in Mejía and Uribe [Nov. 2009] are highly variable and do not display any consistent pattern or trend over the past several years.

The relationship between union homicides and "union activity" (defined below) is then modeled using ordinary least squares regression (OLS) and instrumental variables analyses (IV). The variables used to build these models are described in more detail below. These methods rely on very strong assumptions. We do not believe that Mejía and Uribe [Nov. 2009] adequately addresses the data's potential violations of these assumptions and the potential ramifications on the estimates when these assumptions are violated. We outline each of these assumptions and potential results of their violation in section 4.

Analyses in Mejía and Uribe [Nov. 2009] rely on data collected by the ENS and the CUT on the number of union member and union leader homicides in Colombia for each year and Colombian department (equivalent to a U.S. State). In all models, the outcome of interest is the homicide rate for either union members or union leaders. "Union activity" is divided into two types: Type I - wage agreements and pacts and Type II - acts of protest, such as strikes or work stoppages. We worry that the construction of the variable "union activity" used in Mejía and Uribe [Nov. 2009] ignores important aspects of union activism in Colombia and abroad, such as the practice by union affiliates of filing formal complaints about breaches of labor rights and corruption in the workplace (*denuncias* in Spanish), among others. Furthermore, it is worth noting that both Type I and Type II activities are not exclusive activities of union affiliates. Other types of workers and activists may lead or participate.

The models in Mejía and Uribe [Nov. 2009] also adjust for GDP per capita, total homicide rate, number of police arrests per 100,000 individuals, number of paramilitary and guerilla attacks on civilians, and fixed effects for year and Colombian department. The IV analysis uses two instruments, one for each type of union activity. An instrument is a variable that is believed to be related to the outcome of interest (union homicide rate) only through its correlation with the explanatory

[1] Convenience samples are collected without the use of an underlying random selection method. Instead, they are based on available, observable data. As such, they are useful for documentation purposes, but they cannot be relied upon to represent an underlying population or to accurately describe patterns over time and space. See Davenport and Ball [2002], Guberek et al. [2010] and Gohdes [2010] for more details about ways in which conclusions based on convenience samples can be biased.

variable of interest (union activity) (see Wooldridge 2002 for a more formal definition). The instrument for Type I activity is "... the degree of formality of labor markets in the industry [as measured by] the percentage of full time employees with open-ended contracts and social security payments per capita." (p.11) The instrument for Type II activity is industrial activity, as measured by "per capita industrial consumption of energy and the number of industry establishments per capita." (p.11)

The paper concludes that since the measures of union activity are not significant in the majority of the models, there is no relationship between union homicide rates and union activity, and thus that "... violence against unionists in Colombia is neither systematic nor targeted" (p.1) One of the main modeling assumptions which we believe these data violate, uncorrelated errors, has a direct affect on significance tests. Therefore we are highly skeptical about the reliability of the conclusion that union activity and union violence are not significantly associated.

It is worth noting that Mejía and Uribe [Dec. 2009] presents additional tables not in the initial paper [Mejía and Uribe, Nov. 2009]. These tables replicate the analyses in the original paper using a new truncated outcome variable (union member homicide rate) that excludes the lower and upper 5% of observations. This supplemental material is used to build the argument that since similar conclusions can be reached based on an outcome that excludes potential outliers, the high level of variability in the union member homicide rate is not affecting the ability to detect a relationship between this outcome and union activity. We disagree with this interpretation of the supplemental tables - indeed we do not believe that they result in similar conclusions to those presented in the initial paper. Instead, we find these supplemental tables to be further evidence of the large amount of variability present in counts of union homicides, inconsistent modeling results that can occur from relatively small changes in the data and the general lack of reliability of conclusions based on these models. More details of this comparison are available in the section "Interpretation of Descriptive Statistics and Results" below.

4 Questions and Concerns Regarding Data and Methodological Choices

The paper presented by Mejía and Uribe raised many questions and concerns for us. Specifically, although Mejía and Uribe [Nov. 2009] includes a summary of the data, a broader reflection about the scope and limitations of these data is lacking. We are particularly concerned about basing such strong causal conclusions on convenience sample data, which are not designed to support statistical inference, much less causal inference. Additionally, it is not clear to us from the paper how the strong assumptions of the techniques presented in Mejía and Uribe [Nov. 2009] were met. Violations of these assumptions could change the magnitude of parameter estimates (used to quantify the relationship between union activity and union violence) and the significance of parameter estimates (used to determine the presence or absence of any relationship).

We have organized these sections in decreasing magnitude of the potential ramifications of our concerns regarding the data and methodology.

Under-Registration. As mentioned in Section 3 above, the data in Mejía and Uribe [Nov.

2009] come from convenience samples. Since convenience samples rely on available, observable data, the patterns detected in these kinds of samples are *reporting* patterns, which may or may not reflect underlying patterns of violence in the larger population. For example, in a particularly violent region, keeping a record of union member homicides might be a dangerous activity. Therefore no source records union member homicides in that particular region. Analyses based on these data would conclude that there were low levels of violence in that region, whereas the truth would be that there were low levels of *reports* of violence in that region. This results in under-registration - a lack of recorded events where in fact events do exist. See Davenport and Ball [2002], Guberek et al. [2010] and Gohdes [2010] for further details about and examples of under-registration and reporting bias. Although Mejía and Uribe [Nov. 2009] does include several robustness tests confirming that the presented conclusions are not sensitive to different data sources, this does not change the fact that all data sources may be sensitive to similar reporting biases, simply by the nature of the data being collected.

This is a common challenge with any dataset that is considered a convenience sample [Ball, 2000, AAAS, 2002, Ball et al., 2003, Silva and Ball, 2006]. That is, any data that is not collected using an underlying random process, but rather records available information. Despite best efforts, groups collecting data in this way may have different levels of access, available resources, or security situations in different geographic regions or over time. For these reasons it is normal, even expected, that convenience samples do not include, or represent, the entire universe under study. The problem is that we do not know the structure of the missingness - we cannot describe which pieces of the universe a convenience sample is missing.

For example, although Figure 4 (a) [Mejía and Uribe, Nov. 2009, p.22] shows that data from the Office of the Vice President and ENS have very similar patterns, it also reveals that those two datasets do not report identical homicide rates (ENS consistently reports higher rates than the Office of the Vice President). If the Office of the Vice president is missing some of the observations recorded by ENS, how can we know with certainty that ENS is not also missing some observations? What will be the impact of this?

Additionally, we do not know that the data from the Office of the Vice President are a perfect subset of the ENS data. In other words, both datasets may be reporting similar patterns based on different observations. Given the well-documented limitations of convenience samples [Dorofeev and Grant, 2006] and the fact that we have no way of quantifying the under-registration, and thus of quantifying the resulting potential bias, we would be quite hesitant to base such strong conclusions about union violence on these data.

Lastly, all types of convenience data, not just data on violence, are susceptible to under-registration. Therefore, measures of union activity may also be incomplete and fail to represent the underlying population of union members and their activities. Specifically, as mentioned in Section 3, the definition of union activity used in Mejía and Uribe [Nov. 2009] omits an important aspect of union activity in Colombia and includes activities in which any worker, not just union members, may participate. For these reasons we suspect that the measure of union activity used in the analyses in Mejía and Uribe [Nov. 2009] may not accurately represent the underlying population of union activity and may therefore contribute to biased estimates.

Panel Data and Autocorrelation. The empirical exercise "...to test the hypothesis that greater union activity causes more homicides of union members and union leaders," (p.8) in Mejía and Uribe [Nov. 2009] relies on what is called *panel data*. This is data with both spatial and temporal dimensions. More specifically, the variables described in Section 3 were collected in each Colombian department for each year between 2000 and 2008. This kind of data requires "...special treatment because of the correlation across time." [Wooldridge, 2002, p.18] For example, it is plausible that the quantity of reported homicides in a specific Colombian department and year could affect the quantity reported in the same department in another year. If in one year a large number of union members are killed, then the following year perhaps fewer homicides would be reported. This is because groups recording homicide incidents would consider that a high risk activity and reduce their presence in that region. Alternatively, following a particularly violent year, perhaps fewer individuals would choose to be union members or participate in union activities. Regardless of the reason, the reporting rates of homicides of union members in those two years would be correlated. Technically, this problem is referred to as autocorrelation.[Davis, 2002]

Although the model presented in Mejía and Uribe [Nov. 2009] includes subscripts for Colombian department[2] and time

$$HRUM_{s,t} = c_1 + \gamma UA_{s,t} + \beta X_{s,t} + e_{s,t} \tag{1}$$

(Eq 1, p.10; $HRUM_{s,t}$ = Homicide Rate for Union Members in Department s at time t, c_1 = constant, $UA_{s,t}$ = Union Activity in Department s at time t, $X_{s,t}$ = all other explanatory variables, $e_{s,t}$ = error term), we were unable find any specific references or details as to how the autocorrelation problem was addressed. The identification of autocorrelation, and choice of which specific autocorrelation structure[3] to incorporate into an analytical model, can have a large impact on the results estimated by that model. More specifically, if autocorrelation is ignored, and standard errors are not calculated such that they are robust to such correlation, then significance tests (which rely on standard errors) may be incorrect.[Wooldridge, 2002] The strong conclusions in Mejía and Uribe [Nov. 2009] are based entirely on the lack of a significant association between union activity and union homicide in these models. If these significance tests are incorrect due to the presence of autocorrelation, we must reject these conclusions.

Alternatively, if a default correlation structure was used, without evaluating whether it is the most appropriate structure for this specific dataset, the conclusions presented in this paper may be based on a weaker model than could potentially be identified with the proper correlation structure. Lastly, it is possible that a specific correlation structure was defined and not explicitly described in the methodological section of the paper. If this is the case, we still have questions about how a correlation structure was selected and how correlation parameters were estimated.

Instrumental Variables (IV). The potential endogeneity problem resulting from a model such as Equation 1 is correctly identified in Mejía and Uribe [Nov. 2009]. That is, "...the intensity of union activity ($UA_{s,t}$) is an endogenous variable, since it could be affected by the degree of violence against union members." [Mejía and Uribe, Nov. 2009, p.10] An instrumental variables approach is commonly used in economics to solve the problem of endogeneity. However, we want to

[2]Colombian department is noted in the equations with s, as used in Mejía and Uribe [Nov. 2009]

[3]Potential correlation structures include independent, pairwise independent, exchangeable, and unstructured.[Agresti, 2002, Johnson and Wichern, 1998]

highlight that not all IV assumptions are testable (Wooldridge 2002, Gelman and Hill 2006, among others).

In particular, the IV model (Equation 2 Mejía and Uribe Nov. 2009, p.11)

$$UA_{s,t} = c_2 + \delta_1 z_{1s,t} + \delta_2 z_{2s,t} + \beta X_{s,t} + u_{s,t} \tag{2}$$

relies on the assumptions that $corr(z_{1s,t}, e_{s,t}) = 0$ and $corr(z_{2s,t}, e_{s,t}) = 0$ (where $z_{1s,t}$ and $z_{2s,t}$ are the instruments described in Section 3 and $e_{s,t}$ is the error term from Equation 1). In other words, that the instruments are uncorrelated with the error term. These assumptions cannot be tested. Instead, they rely on "economic behavior or a gut feeling." [Wooldridge, 2002, p.463] Other analysts must share these *opinions* about the relationships between union activity, formality of labor markets, industrial activity, and union member homicide rates to have confidence in the conclusions put forth in Mejía and Uribe [Nov. 2009]. Although the formality of labor markets and industrial activity appear to be reasonable instruments, other equally reasonable instruments are also likely to exist. For example, the timing of a legal change, such as a law that might affect labor market formality, might be a suitable alternative instrument.

We find the instrumental variable analysis carried out in Mejía and Uribe [Nov. 2009] to be methodologically correct, however we do have one clarifying question: In Equations 1 and 2 on pages 10 and 11 respectively, how can the same parameter (β) describe the relationship between X and $HRUM$ and X and UA?

Parameter estimation. As mentioned in the previous section on instrumental variables, although we find the analyses in Mejía and Uribe [Nov. 2009] to be methodologically correct, they do not necessarily reflect methodological choices that we would make. We find exploring the motivations behind methodological choices, and comparing potentially different results and conclusions based on alternative methods, to be beneficial in understanding complex relationships, such as the one between union activity and union violence. We are specifically interested in further exploring both the method used in Mejía and Uribe [Nov. 2009] to estimate the model parameters described above and the specific model parameter on which conclusions are based.

Modeling and parameter estimation is a broad category within statistics (and econometrics) and there exist many specific methods within this broad category. Mejía and Uribe [Nov. 2009] uses two of the most popular methods to estimate model parameters, ordinary least squares (OLS) and two stage least squares (2SLS). Given the autocorrelation problem described above, OLS is arguably not the best method to estimate the parameters. As mentioned in Plümper et al [2005] "[t]here are four potential violations of OLS standard assumptions in panel data: errors tend to be autocorrelated (serial correlation of errors) - that is, they are not independent from one time period to the other . . . ". There are alternative methods that address this problem, and estimate the model parameters appropriately in the presence of autocorrelation. For example, Beck and Katz [1996] recommend adding a time-lagged *dependent* variable to the right-hand side of the modeling equation (Mejía and Uribe [Nov. 2009] includes a model with a time lagged *independent* variable to address the reverse-causality problem). We would be very curious to see what affect such an analysis had on the magnitude and direction of current parameter estimates.

Secondly, the parameter of interest in Mejía and Uribe [Nov. 2009] is γ in the model:

$$HRUM_{s,t} = c_1 + \gamma UA_{1s,t} + \beta X_{s,t} + e_{s,t}.$$

In this model γ represents the mean effect of $UA_{s,t}$ over $HRUM_{s,t}$. Given the potential complexity in the relationship between $UA_{s,t}$ and $HRUM_{s,t}$ it seems to us that it would be more interesting to have a set of parameters γ_s, which provide the relationship between these two variables for each Colombian department. This is possible using a model with random effects instead of fixed effects. This would require a larger sample size, and perhaps this analysis is not possible using the specific data from Mejía and Uribe [Nov. 2009]. However, we think this is a more appropriate analysis to address the question of interest posed in Mejía and Uribe [Nov. 2009].

Interpretation of Descriptive Statistics and Results. At several points in Mejía and Uribe [Nov. 2009] we find ourselves disagreeing with the interpretations put forth about both descriptive graphs and model results. Specifically, Figures 3, 4(a) and 4(b) are described as steadily decreasing or showing a sustained reduction (p. 5-6). However we interpret all three figures as showing inconsistent or relatively flat patterns. In particular, the pattern in Figure 4 (b) for the years 2003-2008 clearly alternates between an increasing and decreasing ratio. We would not be comfortable describing any of these graphs as depicting a steadily decreasing trend. Although these descriptive graphs are not the primary evidence for the conclusions in Mejía and Uribe [Nov. 2009], we still feel it is vital to describe the data accurately at every stage of analysis.

We also consider these descriptive graphs to indicate just how highly variable the union member homicide rate is. Indeed, in Table 1 [Mejía and Uribe, Nov. 2009, p.24] the three reported homicide rates (for union members, leaders, or workers) range from a low mean of 7.7 (per 100,000) with a standard deviation of 37 to a high mean of 25.1 (per 100,000) with a standard deviation of 38.1. In addition to the challenges posed by the structure of panel data, an outcome variable with a large variance can be quite challenging to model. This fact, coupled with our skepticism about basing statistical inference on convenience data, makes us all the more hesitant to base strong quantitative conclusions on this analysis.

Turning now to the empirical strategy, Mejía and Uribe [Dec. 2009] argues that other variables included in the model (such as the total homicide rate and GDP per capita) are consistently significantly associated with the union homicide rate. This is stated as evidence that the high level of variability in the measure of union homicide rates does not negatively affect the estimation of model parameters. However, looking across all of the tables, we find that few control variables are consistently significantly associated with union member homicide rate.

In table 2a, the OLS analysis, it does appear that total homicide rate and police arrests are consistently associated with union member homicide rate. However, the *direction* of the association between union member homicide rate and total homicide rate switches in the final model (positive parameter estimates in models 1-7, negative in model 8). Also in this analysis, GDP per capita is significant in models 5 and 7, not significant in models 3, 4, 6, and 8, and, similar to total homicide rate, the *direction* of the association changes (from negative to positive) in the final model.

In table 2b, simply adding a lag time to the union activity variable changes some of the relationships reported in table 2a - in particular, models 5-7 indicate a non-significant relationship

between total homicide rate and union member homicide rate before this term becomes significant again in the final model.

In table 2c, the first IV model, total homicide rate remains significant until the final model, while GDP per capita flips between significant (models 3, 4, and 6) and not significant (models 5, 7, and 8). Additionally, the final model (8) appears to explain 94% of the variation in union homicide rates (R-squared = 0.944) yet none of the variables in this model are significant. Perhaps most interestingly, when this table is reproduced in Mejía and Uribe's supplemental material [Dec. 2009] using a version of union member homicide rate that excludes potential outliers, these results change drastically. None of the other explanatory variables are significant in any of the models, and union activity is significant in the first four models. This seems to contradict both the conclusion that union activity is not significantly associated with union homicide rate and that results are robust to potential outliers in the union homicide rate.

We believe that close examination of the tables presented by Mejía and Uribe in both their original paper and supplemental materials indicate erratic model behavior. Such inconsistent model results likely stem from an outcome variable that is very difficult to explain using this modeling method. Again, this would lead us to draw more modest conclusions based on these data and these methods.

Distribution of HRUM. Lastly, we did not find in the paper the distribution assumption associated with the outcome variable (homicide rate of union members - $HRUM$). We assumed that $HRUM \sim N(\mu, \sigma^2)$ was used, however explicit mention of this modeling assumption, and the justification for this choice, would better enable us (and others) to evaluate this analysis. Additionally, a preferable distribution for modeling rates, particularly those such as $HRUM$ which can typically be expected to be asymmetric, is $HRUM \sim Beta(\alpha, \beta)$ [Ferrari and Cribari-Neto, 2004]. Since assumptions about the distribution of the outcome variable drive modeling choices, assuming a normal distribution when the underlying data are in fact skewed could potentially lead to bias in the estimated regression coefficients and misleading conclusions.

5 Conclusions

We believe that the findings in Mejía and Uribe [Nov. 2009] are overstated. We find that weaknesses in the data, in the model choice, and in the model interpretation all raise serious questions about the strong causal conclusions presented in Mejía and Uribe [Nov. 2009]. Specifically, we have the following concerns:

The descriptive and analytic results indicate that union member homicide rate is a highly variable outcome measure. In some times and places, there are zero homicides, whereas in other times and places, there are many. This variation results in poor-quality models. We find that control variables in the various model formulations show inconsistent and indeed reversed effects, suggesting problems with the model specification, with the data on homicides, or both.

Given the source (convenience sample) and structure (panel) of the data, we are not convinced that the methodological choices made in Mejía and Uribe [Nov. 2009] are appropriate. In specific technical terms, we are concerned that neither the OLS nor IV analyses adequately adjust for the

inherent autocorrelation of the data. The central conclusion in Mejía and Uribe [Nov. 2009] is based on a lack of a significant association between union activity and union violence. However, the significance tests used for this conclusion are precisely the statistical result most likely affected by autocorrelation. In short, the finding that there is no relationship between union activity and union violence may be an artifact of a failure to adjust for autocorrelation.

The poor quality of these models, the unknown under-registration inherent in the data, and the questionable modeling decisions mean that any conclusions should be very carefully qualified. Unfortunately, Mejía and Uribe [Nov. 2009] presents quite strong conclusions, which we believe are unsupported by the analyses.

Finally, as is best scientific and academic practice, Mejía and Uribe have offered to share data and models with peers to engage in open and transparent review. We welcome this offer. Furthermore we would like to underscore the value for historical truth that this methodological debate is taking place openly and transparently.

Acknowledgements

We are grateful to Tamy Guberek, Romesh Silva, and Patrick Ball for their invaluable input editing and structuring this review. We would also like to thank Dylan Small for his rapid review of our work. Many thanks to Beatriz Vejarano for translating this document into Spanish.

References

Alan Agresti. *Categorical Data Analysis*. John Wiley & Sons, Inc., 2002.

Patrick Ball. *The Guatemalan Commission for Historical Clarification: Generating Analytical Reports; Inter-Sample Analysis*, chapter 11. American Association for the Advancement of Science (AAAS) Science and Human Rights Program, Washington, DC, USA, 2000.

Patrick Ball, Jana Asher, David Sulmont, and Daniel Manrique. How many peruvians have died? 2003.

Nathaniel Beck and Jonathan N. Katz. Nuisance vs. substance: Specifying and estimating time-series-cross-section models. *Political Analysis*, 6:1–36, 1996.

Christian Davenport and Patrick Ball. Views to a kill: Exploring the implications of source selection in the case of guatemalan state terror, 1977-1996. *Journal of Conflict Resolution*, 46:427–450, 2002.

Charles S. Davis. *Statistical Methods for the Analysis of Repeated Measurements*. Springer, 2002.

Sergey Dorofeev and Peter Grant. *Statistics for Real-Life Sample Surveys: Non-Simple-Random Samples and Weighted Data*. Cambridge University Press, 2006.

Silvia L.P. Ferrari and Francisco Cribari-Neto. Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, 2004.

American Association for the Advancement of Science: Science, Human Rights Program, American Bar Association (ABA), Central, and East European Law Initiative. Political killings in kosova/kosovo, march-june 1999: A cooperative report by the central and east european law initiative of the american bar association and the science and human rights program of the american association for the advancement of science. Washington, DC, USA, 2002.

Andrew Gelman and Jennifer Hill. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, 2006.

Anita Gohdes. Different convenience samples, different stories: The case of sierra leone. 2010.

Tamy Guberek, Daniel Guzmán, Megan Price, Kristian Lum, and Patrick Ball. To count the uncounted: An estimation of lethal violence in casanare. 2010.

Richard A. Johnson and Dean W. Wichern. *Applied Multivariate Statistical Analysis*. Prentice-Hall, Inc., 1998.

Daniel Mejía and María José Uribe. Respuesta a los comentarios de la escuela nacional sindical al artículo: "is violence against union members in colombia systematic and targeted?". Dec. 2009.

Daniel Mejía and María José Uribe. Is violence against union members in colombia systematic and targeted? Nov. 2009.

Thomas Plümper, Vera E. Troeger, and Philip Manow. Panel data analysis in comparative politics: Linking method to theory. *European Journal of Political Research*, 44:327–354, 2005.

Romesh Silva and Patrick Ball. *"The Profile of Human Rights Violations in Timor-Leste, 1974-1999." A Report by the Benetech Human Rights Data Analysis Group to the Commission on Reception, Truth and Reconciliation*. 2006. URL `http://www.hrdag.org/timor`.

Jeffrey M. Wooldridge. *Introductory Econometrics: A Modern Approach*. South-Western College, 2002.

About the Benetech Human Rights Program

The Benetech Human Rights Program has more than 17 years of experience applying rigorous scientific analysis to data about human rights violations. Our expertise has been sought by nine Truth and Reconciliation Commissions, by U.N. missions and official human rights bodies, and by international criminal tribunals. We have conducted projects in El Salvador, Ethiopia, Guatemala, Haiti, South Africa, Kosovo, Sierra Leone, Sri Lanka, Timor-Leste, Colombia and Peru; and provided extensive guidance on data processing and analysis methodologies to non-governmental organizations and partner groups in many countries throughout the world. With our partners, we make scientifically-defensible arguments based on rigorous evidence.[4]

This project was funded by core support to the Benetech Initiative from the John D. and Catherine T. MacArthur Foundation, the Sigrid Rausing Trust and The Oak Foundation, as well

[4] (`http://www.benetech.org`, `http://www.hrdag.org`).

as project specific funding from the European Commission.

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 or to this project.

The Benetech Initiative
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