

A Re-assessment of Reporting Bias in Event-based Violence Data with Respect to Cell Phone Coverage

Hollenbach, Florian M.; Pierskalla, Jan H.

Document Version
Final published version

Published in:
Research and Politics

DOI:
[10.1177/2053168017730687](https://doi.org/10.1177/2053168017730687)

Publication date:
2017

License
CC BY-NC-ND

Citation for published version (APA):
Hollenbach, F. M., & Pierskalla, J. H. (2017). A Re-assessment of Reporting Bias in Event-based Violence Data with Respect to Cell Phone Coverage. *Research and Politics*, 4(3), 1-5.
<https://doi.org/10.1177/2053168017730687>

[Link to publication in CBS Research Portal](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us (research.lib@cbs.dk) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 10. Oct. 2024



A re-assessment of reporting bias in event-based violence data with respect to cell phone coverage

Research and Politics
July-September 2017: 1–5
© The Author(s) 2017
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/2053168017730687
journals.sagepub.com/home/rap


Florian M. Hollenbach¹ and Jan H. Pierskalla²

Abstract

This paper discusses the issue of possible reporting bias in media-based violent-event data and its relation to the role of communication technology in fostering collective action. We expand the work of Weidmann (2016), presenting several sensitivity analyses to determine the degree to which reporting bias may confound the relationship between communication technology and violence in a recent study that relies on event data for Africa. We find no strong evidence that suggests results on the positive relationship between communication technology and collective action in the study by Pierskalla and Hollenbach (2013) are wholly an artifact of reporting bias.

Keywords

Communication technology, collective action, reporting bias, event data

Introduction

Research into the effects of modern information and communication technology (ICT) on collective action suggests that the spread of cell phone technology can play an important role in facilitating various types of political collective action (Bailard, 2015; Manacorda and Tesei, 2016; Pierskalla and Hollenbach, 2013). For example, using spatially disaggregated violent event data, Pierskalla and Hollenbach (2013) document a positive association between cell phone coverage and violent events on the African continent between 2007 and 2009.

At the same time, these findings could potentially be driven by the presence of reporting bias. This issue is particularly relevant for scholars working with event data that is based on media reports (Croicu and Kreutz, 2016; Weidmann, 2016). Addressing reporting bias is especially crucial if new communication technology increases the ability of news organizations to report political events. Then reporting bias is likely to affect inferences about the relationship between ICTs and the occurrence of violence.

While Croicu and Kreutz (2016) find only a small effect of cell phone technology on the quality of reporting in African violent-event data, a study by Weidmann (2016), using high-quality data from Afghanistan, suggests that findings about the effects of cell phone technology on the incidence of violent events might be mere artifacts of

reporting bias. He shows that there is a positive association between cell phone coverage and violent events in Afghanistan based on data from the Uppsala Conflict Data Program - Georeferenced Event Dataset (UCDP-GED); however, the relationship is not statistically significant (at standard levels) when relying on event data collected by the US military. In addition, Weidmann (2016) implements a sensitivity analysis for the findings in Pierskalla and Hollenbach (2013), which suggests that the original results are also consistent with reporting bias.

In this note we follow Weidmann's lead and present a more extensive set of sensitivity analyses that explore the issue of reporting bias in Pierskalla and Hollenbach (2013) further. First, we argue that several ad hoc choices over model specification have to be made during the sensitivity analysis (e.g. about "high" or "best" fatality estimates or which specification to estimate). These choices affect the conclusions of the sensitivity check. When we consider a

¹Texas A&M University, College Station, TX, USA

²The Ohio State University, Columbus, OH, USA

Corresponding author:

Florian M. Hollenbach, Texas A&M University, 2061 Allen Building, 4348 TAMU, College Station, TX 77843-4348, USA.
Email: fhollenbach@tamu.edu



broader range of sensitivity tests, we find no clear evidence of systematic reporting bias in Pierskalla and Hollenbach (2013), irrespective of the specific change in model specification. Second, we implement a new simulation-based sensitivity analysis. We add random, additional events to regions without cell phone coverage (directly contradicting the main finding) in order to simulate data without reporting bias (following Gallop and Weschle, 2017). Assuming high levels of reporting bias across the UCDP-GED Africa data, we still recover a positive effect of cell phone coverage on violence in approximately more than 99% of simulations.

In total, these tests reveal that reporting bias is unlikely to completely explain the positive association between ICTs and violent collective action in the context of the Pierskalla and Hollenbach's (2013) study, adding to the ongoing debate on the relationship between modern communication technology and collective action. Finally, this empirical exercise also illustrates different practical strategies for exploring the issue of reporting (or other forms of measurement) bias in applied settings.

Window analysis

Pierskalla and Hollenbach (2013) study the effect of cell phone coverage on violent events in Africa. They match data on the spatial extent of 2G cell phone networks to a set of grid cells, covering the years from 2007 to 2009. The authors estimate the effect of cell phone coverage on the probability of a violent event taking place in each grid cell, relying on UCDP-GED event data (Sundberg and Melander, 2013), and find a robust and positive effect of cell phone coverage on violence.

Weidmann (2016) implements a sensitivity analysis for this finding which rests on a strategy proposed by Dafoe and Lyall (2015). The approach assumes that the magnitude of reporting bias varies depending on the severity of the reported events. Dividing the data into different subsamples, based on the reported fatalities, subsamples with low fatality counts ought to suffer from more reporting bias and vice versa. If bias is driving the finding, then with increasing fatality rates, the estimated coefficient for cell phone coverage should decrease. To create the sensitivity analysis, Weidmann (2016) orders all violent events for 2008 by the number of fatalities. He then creates sub-sets of the data (windows) that include 50% of the violent events in 2008 (444 events), changing the composition of the violence data by sliding the window by 10 events per subset. Weidmann (2016) shows that the marginal effect of cell phones decreases for samples with more fatalities, and the 95% confidence interval covers zero for about one-third of the estimated models, which is consistent with reporting bias being an alternative explanation for the original finding.

This is a useful and straightforward approach, but requires the analyst to make a number of ad hoc model specification choices that might affect the conclusions of the sensitivity check.

For instance, one choice is related to the fatality estimate. The UCDP-GED data includes different estimates for the number of fatalities for each violent event. The data set provides a *best* estimate and a *high* estimate. To order the events by severity, Weidmann (2016) uses the *high* estimate. There is no clear reason to prefer the *high* over the *best* estimate and vice versa. A second matter is the model specification for the windowed analysis. Weidmann (2016) presents the results from the most basic model in Pierskalla and Hollenbach (2013). Here, we also present the windowed analysis for the panel model, a preferred and more conservative specification (more details can be found in the Supplementary Appendix).

Figure 1 replicates Weidmann's sensitivity check in panel (a) and switches to the "best" fatality estimate in panel (b). Panel (c) shows coefficient estimates for the panel data using the "high" estimate, and panel (d) for the "best" estimate. Several things stand out. For panel (b) the coefficient size first increases and then decreases. In panel (c) and (d) coefficient estimates are relatively stable across windows, except for very large events where the effect is actually stronger. Confidence intervals also do not cover zero for the vast majority of windows. None of these patterns are consistent with reporting bias as in panel (a).

Naturally, researchers could make a number of other alternative choices about model specification. In Section 1 and 2 of the Supplementary Appendix we consider several of them, largely following the robustness checks in the original study by Pierskalla and Hollenbach (2013). For example, we present windowed analyses that include country means, country fixed effects, and a spatial lag for the cross-sectional model. For the panel data we also consider event counts as dependent variables, precisely coded event data, and event data based on a newer version of the UCDP-GED data. All in all, the results are rarely indicative of reporting bias (even with windows created on the *high* fatality estimates).

Simulation analysis

Building on Gallop and Weschle (2017), we simulate data to account for the presence of reporting bias in the event data. If reporting bias is of concern, then violent events are especially under-reported in areas without cell phone coverage. In our cross-sectional data for 2008 there are 358 grid cells with at least one event; 182 of those are in areas without cell phone coverage. Similarly, our three-year panel data includes a total of 922 grid-cell years with at least one violent event; 529 of those are in areas without cell phone coverage. If reporting bias is present, 182 and 529 represent a lower bound on the actual number of events in areas without coverage.

Therefore, we randomly add additional events to grid cells without cell phone coverage. Doing so generates data patterns that go directly against the main finding and will

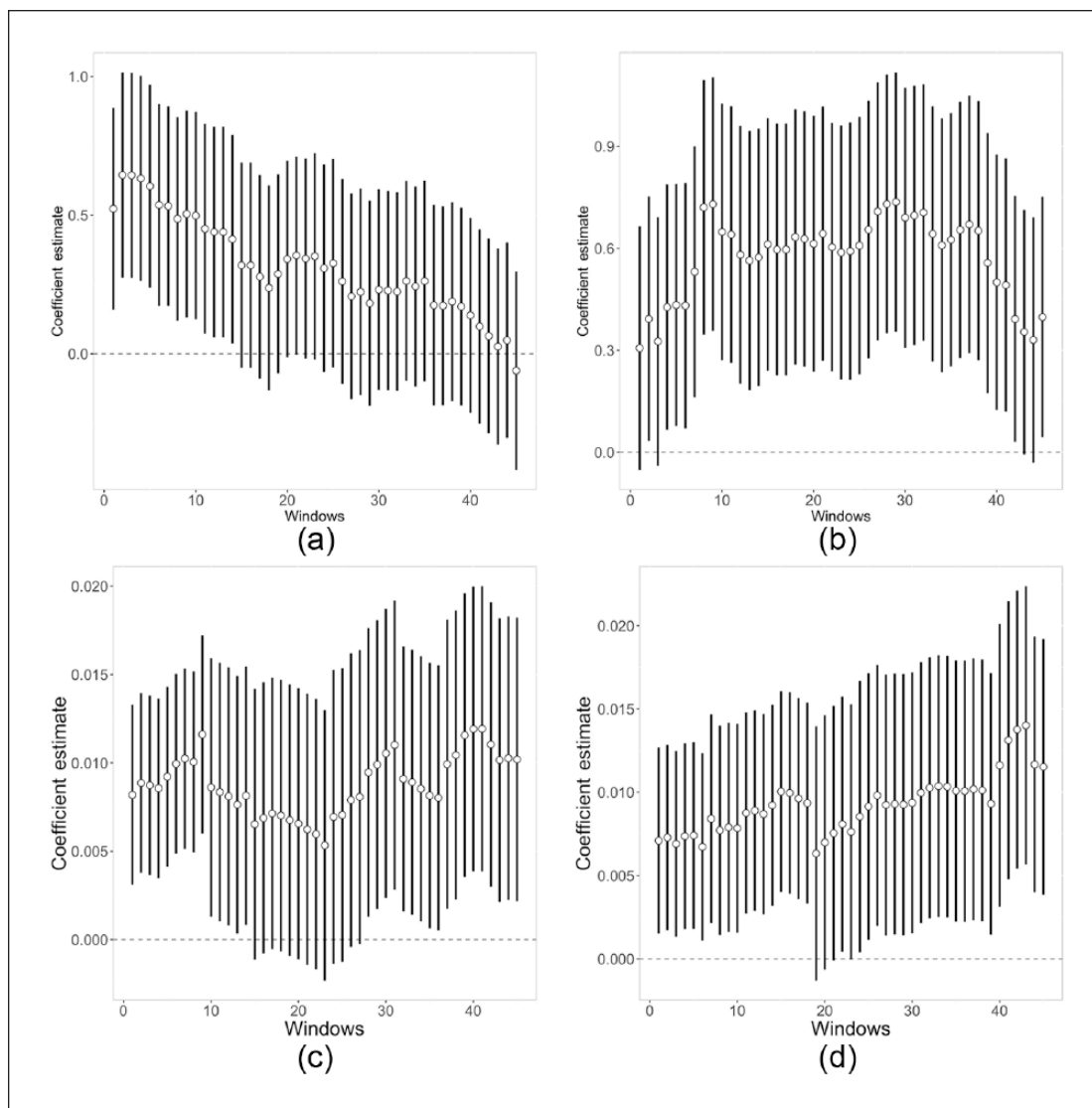


Figure 1. Coefficient estimates for the windowed simple logit using the “high” fatality estimate in panel (a) and the “best” fatality estimate estimate in panel (b). Panel (c) and (d) show the same for the panel data. Only the first plot is consistent with reporting bias.

eventually push the estimated coefficient for cell phone coverage toward zero. First, we estimate a cross-sectional model without the cell phone coverage variable. We then use the predicted probabilities from the model without cell phone coverage to draw “fake” observations of violent events in grid cells where no violence was reported and cell phone coverage is not existent.¹ We simulate new data for five different scenarios, assuming that the observed data in the non-cell coverage areas represent 95%, 90%, 85%, 80%, and 75% of “real” events. For each of these potential levels of reporting bias we create 1000 data sets with additional events and estimate the preferred statistical model. Using the estimated coefficients and standard errors from each simulation, we draw 200 coefficients from a multivariate normal distribution. We then combine all 20,000 draws to generate the overall coefficient distribution under

each scenario. This empirical distribution reflects variation across the simulation runs and integrates the estimation uncertainty expressed in the standard errors. Figure 2 shows the empirical distribution of these coefficients for the cross-sectional linear probability model with country fixed effects.

Across panels (a)–(e) in Figure 2 we observe that the estimated coefficient is nearly always positive; that is, even after correcting for substantial reporting bias, we would still find evidence supporting the original finding.

Without information on the true size of reporting bias, it is hard to know exactly whether 5%, 10%, or 25% of underreporting is the most likely scenario. In addition, the true effect of cell phone coverage is likely to vary across African countries and conflicts, as is reporting bias. Weidmann’s (2016) analysis provides an empirical

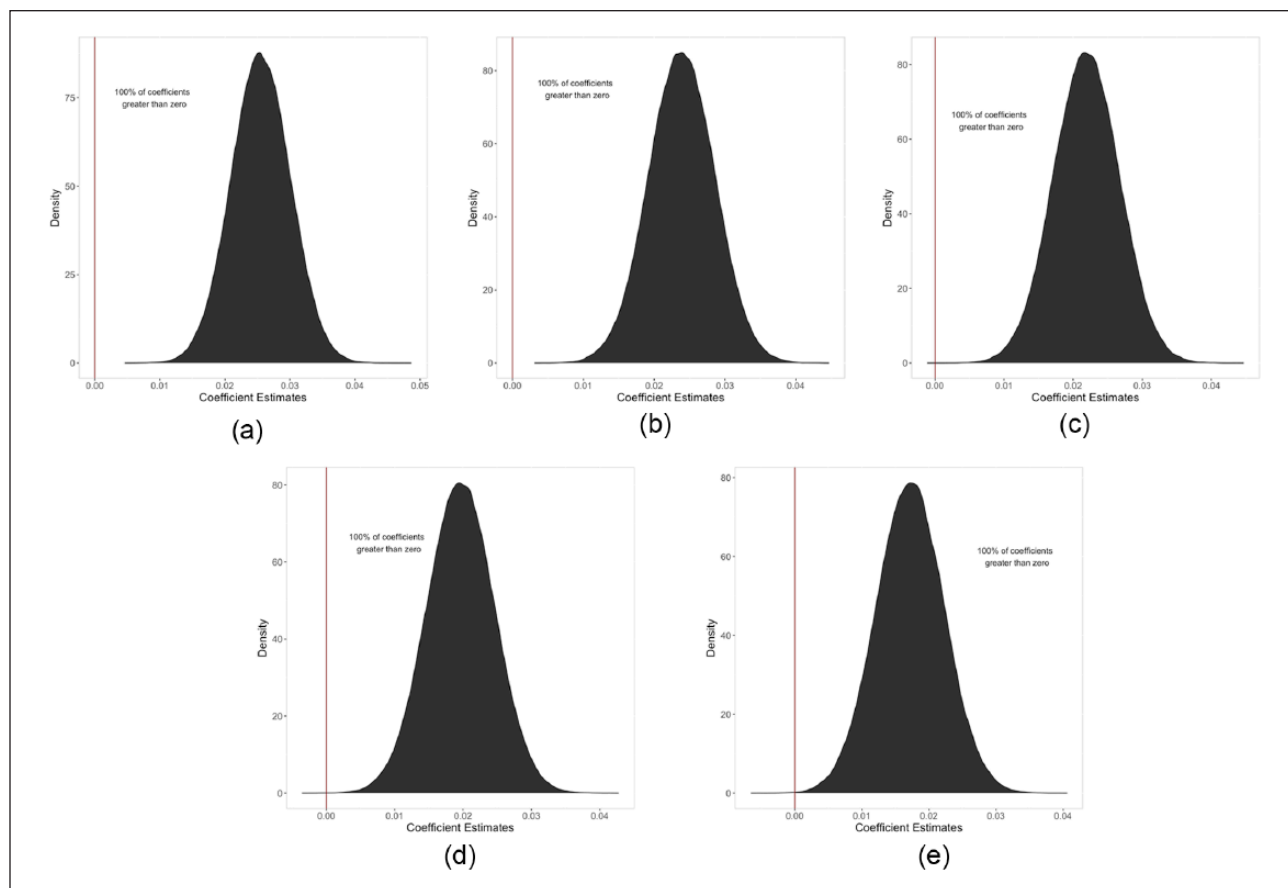


Figure 2. For panel (a) we assume the observed data include 95% of real events. This number decreases by 5 percentage points for each panel. Panel (b) thus represents the results when the observed events are 90% of all true events, panel (c) 85%, panel (d) 80%, and panel (e) assumes that the observed data only include 75% of all real events.

estimate for the case of Afghanistan, indicating that in areas without cell phone coverage, observed events in the data are about 76% of events that would have been observed with cell phone coverage present. This corresponds to the simulation results presented in Figure 2(e), which suggests that the original findings in Pierskalla and Hollenbach (2013) would be robust to this level of reporting bias. The results are more sensitive to reporting bias in the panel models. Yet still, only if reporting bias is very large could it completely explain the positive association found in the panel models (see Section 4 in the Appendix). We believe these simulations suggest that, in general, there is little indication that reporting bias is driving the overall association between cell phone coverage and conflict in the sample of African grid cells.

Conclusion

We agree with Weidmann's (2016) assessment that reporting bias is a serious problem and represents a true challenge, especially when trying to infer the effects of ICTs on violent collective action. The additional sensitivity analyses we

present here suggest that there is no overwhelming indication that reporting bias is completely driving the findings in Pierskalla and Hollenbach (2013). Our additional tests also illustrate the practical difficulties in implementing a sensitivity analysis as suggested by Dafoe and Lyall (2015): researchers have to make several ad hoc decisions when conducting sensitivity checks and, hence, ought to consider a broad set of models. We also provide a practical example, drawing on Gallop and Weschle (2017), of how to integrate the simulation of reporting bias in real data.

Beyond the technical issues of sensitivity analyses, the extent and severity of reporting bias in media-based event data is likely to vary dramatically with context. Similarly, it is likely that the effect of ICTs on political events is heterogeneous across countries. In the absence of "ground truth" data that does not suffer from potential reporting bias, there is no good way to directly determine the extent of reporting bias across contexts. This means it is not straightforward to assess the threats of reporting bias without access to better data and contextual knowledge around the mechanisms producing violent collective action. We hope future studies that rely less on media-based event data will be able to

better delineate when (and how) ICTs can foster violent and non-violent collective action.

Acknowledgements

We thank Nils Weidmann, Simon Weschle, Erik Wibbels, three anonymous reviewers, and the editors of *Research & Politics* for their helpful comments on this manuscript. All remaining errors are ours alone. Authors are listed in alphabetical order; equal authorship is implied. Portions of this research were conducted with high-performance research computing resources provided by Texas A&M University (<http://hprc.tamu.edu>).

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Supplementary materials

The supplementary files are available at: <http://journals.sagepub.com/doi/suppl/10.1177/2053168017730687>. The replication files are available at: <http://thedata.harvard.edu/dvn/dv/researchandpolitics>

Note

1. The covariates included are: violent events before 2000, border distance, capital distance, population in 2005, presence of irrigation systems, mountainous terrain, GDP per capita, and a country dummy. Section 5 in the online Appendix displays the simulation models when “fake” events are added completely at random.

Carnegie Corporation of New York Grant

This publication was made possible (in part) by a grant from Carnegie Corporation of New York. The statements made and views expressed are solely the responsibility of the author.

References

- Bailard CS (2015) Ethnic conflict goes mobile: Mobile technology's effect on the opportunities and motivations for violent collective action. *Journal of Peace Research* 52(3): 323–337.
- Croicu M and Kreutz J (2017) Communication technology and reports on political violence. Cross-national evidence using African events data. *Political Research Quarterly* 70(1): 19–31.
- Dafoe A and Lyall J (2015) From cell phones to conflict? Reflections on the emerging ICT-political conflict research agenda. *Journal of Peace Research* 52(3): 401–413.
- Gallop M and Weschle S (2017) Assessing the impact of non-random measurement error on inference: A sensitivity analysis approach. *Political Science Research and Methods* 1–18. doi:10.1017/psrm.2016.53
- Manacorda M and Tesei A (2016) *Liberation Technology: Mobile Phones and Political Mobilization in Africa*. Working Paper. Available at: <http://www.ieb.ub.edu/files/PapersWSFF2015/WSPEPF2015Tesei.pdf> (accessed 2 June 2016).
- Pierskalla JH and Hollenbach FM (2013) Technology and collective action: The effect of cell phone coverage on political violence in Africa. *American Political Science Review* 107(2): 207–224.
- Sundberg R and Melander E (2013) Introducing the UCDP geo-referenced event dataset. *Journal of Peace Research* 50(4): 523–532.
- Weidmann N (2016) A closer look at reporting bias in conflict event data. *American Journal of Political Science* 60(1): 206–218.