UCD CENTRE FOR ECONOMIC RESEARCH

WORKING PAPER SERIES

2017

David Versus Goliath: Fundamental Patterns and Predictions in Modern Wars and Terrorist Campaigns

Michael Spagat, University of London
Neil Johnson, University of Miami
Stijn van Weezel, University College Dublin

WP17/21

October 2017

UCD SCHOOL OF ECONOMICS
UNIVERSITY COLLEGE DUBLIN
BELFIELD DUBLIN 4

David Versus Goliath: Fundamental Patterns and Predictions in Modern Wars and Terrorist Campaigns *

Michael Spagat

Department of Economics, Royal Holloway University of London

Neil Johnson

Department of Physics, University of Miami

Stijn van Weezel

School of Economics, University College Dublin

Abstract

It is still unknown whether there is some deep structure to modern wars and terrorist campaigns that could allow reliable prediction of future patterns of violent events. Recent war research focuses on size distributions of violent events, with size defined by the number of people killed in each event. Event size distributions within previously available datasets, for both armed conflicts and for global terrorism as a whole, exhibit extraordinary regularities that transcend specifics of time and place. These distributions have been well modelled by a narrow range of power laws that are, in turn, supported by a theory of coalescence and fragmentation of violent groups. We show that the predicted event-size patterns emerge in a mass of new event data covering conflict in Africa and Asia from 1990 to 2014. Moreover, there are similar regularities in the events generated by individual terrorist organizations, 1997-2014. The existence of such robust empirical patterns hints at

*We would like to thank participants in the "Interdisciplinary Perspectives on Modelling Conflict" workshop for many useful suggestions that improved this paper.

the predictability of size distributions of violent events in future wars. We pursue this prospect using split-sample techniques that help us to make useful out-of-sample predictions. Power-law-based prediction systems outperform lognormal-based systems. We conclude that there is indeed evidence from the existing data that fundamental patterns do exist, and that these can allow prediction of future structures in modern wars and terrorist campaigns.

Keywords: Armed conflict, cross-validation, event data, power-law, terrorism
JEL codes: C46, C53, D74

Introduction

Polymath Lewis Fry Richardson showed, in a seminal work, that war sizes follow a fat-tailed distribution which, he suggested, could be well captured by a power law (Richardson, 1960). Later research has updated and confirmed this finding using more rigorous statistical methods (Cederman, 2003; Clauset, 2017). It turns out that the Richardson insight for sizes of whole wars extends to event sizes *within* wars. For this analysis the size of a discrete event, such as a suicide bombing or a battle, is defined by the number of people killed in the event. The distributions of event sizes within nine modern wars are all well approximated by a power law with the estimated power coefficients clustering around 2.5 (Bohorquez et al., 2009). The size distribution for global terrorist events, merging together all events perpetrated by all terrorist groups, is also well captured by a power law with a coefficient around 2.5 (Clauset et al., 2007). This latter finding has practical utility because the identified empirical regularities can be used to predict the probability of a terrorist attack comparable in scale to the 9/11 one (Clauset and Woodard, 2013a,b).

A theoretical conflict model driven by processes of coalescence and fragmentation of groups within warring organizations generates power-law distributions for violent events in which the theoretically derived power coefficients cluster around 2.5 (Bohorquez et al., 2009). Recent extensions and elaborations of this coalescence-fragmentation framework confirm the robustness of the tendency toward 2.5 while also providing further theory that can explain power coefficients going as low as 1.9 and as high as 4.5 (Johnson et al., 2013). It is difficult to observe the inner workings of necessarily secretive insurgencies and terrorist organizations yet there is direct evidence that online

ISIS communities display the coalescence and fragmentation behaviours that are central to the coalescence-fragmentation model (Johnson et al., 2016).

The present paper has four main objectives. First, we exploit a mass of new event data on armed conflict and terrorism (Sundberg and Melander, 2013; National Consortium for the Study of Terrorism and Responses to Terrorism (START), 2016) to offer the most complete exploration ever presented of the empirical patterns in the size distributions of violent events within the contexts of both armed conflict and terrorist campaigns. For our war analysis we use the new version of the data employed in previous research (Johnson et al., 2013), enabling us to extend our reach to no fewer than 202 armed conflicts, including more than 100 Asian conflicts never before included in this research program. Our empirical work on terrorism innovates by operating at the *organization* level, enabling us to demonstrate that the size distributions of violent events perpetrated by 57 individual terrorist organizations resemble the size distributions we find for belligerent groups entangled in armed conflicts. This finding deepens an already identified link between terrorist and insurgent organizations (Bohorquez et al., 2009), which is reassuring given the notorious difficulty in separating the two types of organizations conceptually (Moghadam et al., 2014). Indeed, although it may be possible to draw valid distinctions between insurgent versus terrorist organizations, e.g., concerning their ideologies, they both remain collections of decentralized operatives that must adapt quickly to avoid detection and annihilation. These common pressures should force both types of groups into common David-versus-Goliath tactics that should tend to yield similar attacking patterns and, indeed, we find this in our empirical work.

Our second objective is to evaluate the potential for the family of coalescence-fragmentation models to cover the full range of empirical event-size distributions present in the event data for all of the conflicts and terrorist campaigns we have to work with. We find that these models do perform well because, as they predict, the many new estimated power-law coefficients cluster around 2.5, both for conflicts and for individual terrorist organizations. Nevertheless, we identify a need for further theory that can handle coefficients between 1.5 and 1.9 that appear more than a few times in the data.

Our third objective is to exploit the regularities in the empirical event patterns to make useful predictions about the mixtures of event sizes in future wars, based on the range of empirical power-law coefficients we observe in the data. For this project we keep score on the success rates for our predictions and conclude that they are, indeed, useful. This predictability indicates that we are developing good knowledge of a deep structure of modern wars.

Fourth, we test the predictive performance of the power-law model of armed conflict events against that of a lognormal-based system, the most obvious fat-tailed rival distribution, and find that power-law based prediction systems outperform lognormal ones.

Materials and methods

We take our armed-conflict data from the Georeferenced Event Dataset (GED) of the Uppsala Conflict Data Programme (Sundberg and Melander, 2013). This is the most comprehensive and accurate georeferenced dataset on armed conflict available that systematically collects information on the number of people killed in each event (Eck, 2012; Weidmann, 2013, 2015). The GED records details that include the location, timing, and severity of conflict events along with information on the warring parties that generate these events. The data collection effort covers conflicts between governments and rebel groups, non-state based conflicts (also known as communal violence), and violence perpetrated by the state or insurgency groups against civilians. We use version 4 of the dataset which covers all conflicts in Africa and Asia between 1989-2014. Although the GED offers a global dataset, conflicts in Asia and Africa are covered better than those in Europe and Latin America for which the coverage only extends back to 2005, thereby missing the Yugoslav Wars and much of the conflict in Colombia. The Syrian civil war is currently not included in the GED data. The GED coding rules exclude some low-intensity conflicts by imposing a minimum fatality threshold of 25-battle related deaths in a given year. However, this restriction hardly matters for us since it excludes only minor conflicts that may have been excluded anyway due to not having enough events to allow us to reliably fit a power law to the size distribution their violent events.

We include only true single events in our analysis, removing a small number of aggregate fatality counts that are not broken down to the event level. We also drop conflicts with fewer than 30 events. These screens leave us with 98 African conflicts with 21,239 events and 104 Asian conflicts with 60,162 events. For Afghanistan we split the state-based data into two separate conflicts so that the fighting after the beginning of Operation Enduring Freedom is treated separately from the pre-invasion conflict.

We offer a parallel analysis of terrorist incidents since, as noted above, there is evidence suggesting that terrorist organizations may behave similarly to insurgent groups (Bohorquez et al., 2009; Hammes, 2006; Robb, 2007). For this work we use the Global Terrorism Database (GTD), which is provided by the National Consortium for the Study of Terrorism and Responses to Terrorism (START). A novel feature of the GTD dataset is that it includes both domestic and trans/inter-national terrorist incidents. The GTD is updated annually and provides the most comprehensive dataset on terrorist events that is publicly available. The GTD covers the period from 1970 to 2015 and includes detailed information on incident times, locations, fatality counts and, when identifiable, the perpetrating group or individual. We include only events that are definitely acts of terrorism according to the coding, that are attributed to a known organization and that caused at least one fatality. Finally, we use only events occurring after 1997 because the GTD coding procedures changed in that year. This leaves us with 13,859 terrorist attacks carried out by 57 groups between 1998-2015.

Results

We use the "poweRlaw" package in R (Gillespie, 2014) to fit the model $Ms^{-\alpha}$ to the data for each conflict above an estimated cut-off value s_{min} using maximum likelihood estimation (Clauset et al., 2009; Johnson et al., 2013) where s denotes the number of fatalities in an event, α is the power-law coefficient and M is a normalisation factor ensuring that the cumulative probability distribution sums to unity. Figure 1 provides an example one such fit, this one for the Palestinian-Israeli conflict. Figure 2 plots the estimated α values for the African and Asian conflicts against the p-values of bootstrapped tests of the hypotheses that their data are generated by the fitted power laws for these conflicts. To be clear, each data point in figure 2 summarizes a power law fit for a particular conflict such as the one in figure 1.

The reported p-values are based on bootstrap resampling using 1000 iterations. Most conflicts do have size distributions for their violent events that are well fit by power laws with coefficients clustering around 2.5. At the same time, some conflicts do display α values far from 2.5. Moreover, some conflicts have very low p-values, thereby deviating from the empirically and theoreti-

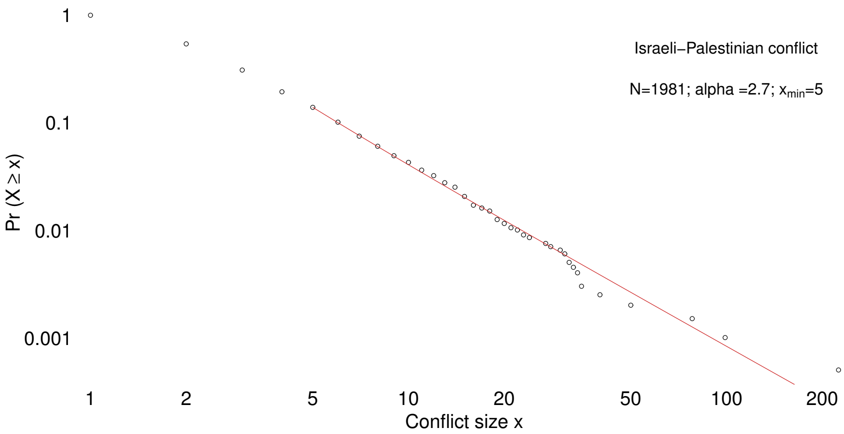

Figure 1: A power law fit for the violent events in the Israeli Palestinian conflict. Logged event sizes on the X axis are plotted against logged probabilities that events are at least as big these sizes.

cally grounded patterns uncovered in earlier research (Bohorquez et al., 2009; Clauset and Woodard, 2013a,b) by suggesting that the power-law hypothesis should be rejected. Low p-values are not necessarily a serious worry since no distribution of violent conflict events will be, literally, generated by an exact power law so we would normally expect to reject the power-law hypothesis with enough data even when this distribution is still useful for modelling the event-generating process of a conflict. Estimated α's far from 2.5, on the other hand, are a more important challenge to the received wisdom in the field. These results could stem from data problems, e.g., not having enough data or having serious flaws in the data-gathering processes for particular conflicts. In fact, in earlier research (Johnson et al., 2013) the conflict in Angola had a very high value of α but now, with a few more years' worth of data, Angola's α has settled in right at 2.5. Or it could be that some modern conflicts really are fundamentally different from the great majority of conflicts we have encountered so far in this research program.

Figure 3 provides the same sort of p versus α information given in figure 2 but this time for terrorist groups using the GTD data. Note that the nature of these results is substantially different from earlier work fitting power laws to global terrorist events (Clauset et al., 2007) because we fit a separate power law to each terrorist organization whereas the previous work merged together all the events generated by all terrorist organizations. It shows that power

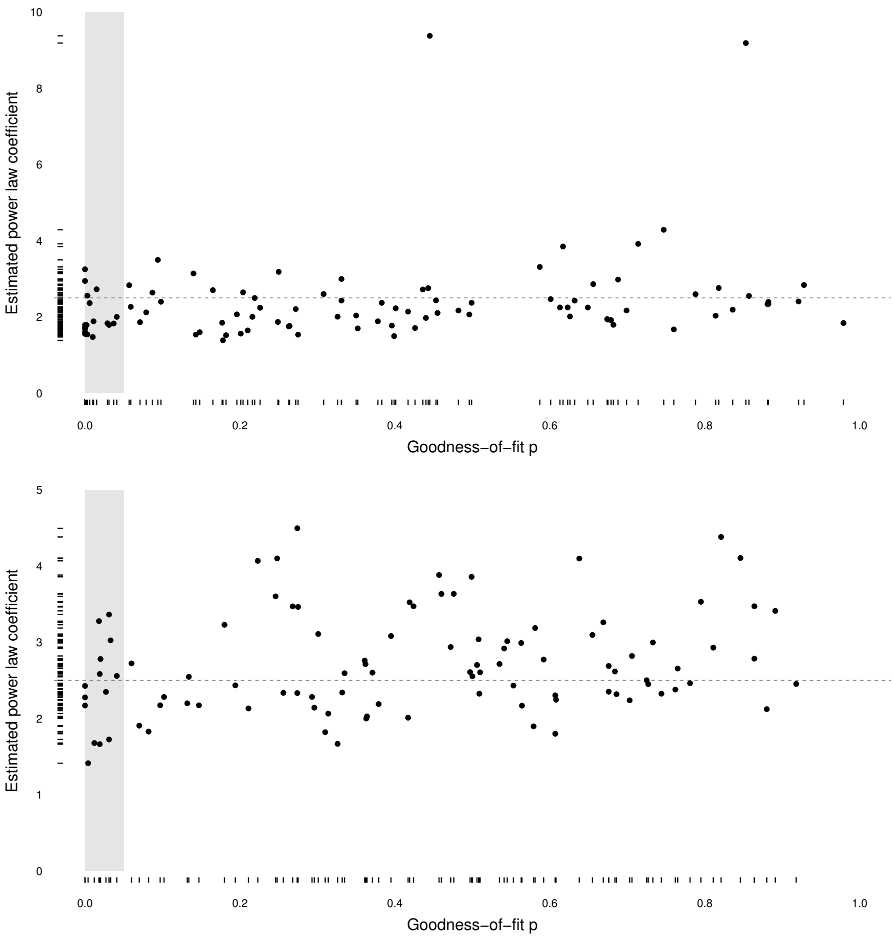

Figure 2: Estimates of α parameters versus p values for power-law hypotheses for African (*top*) and Asian (*bottom*) conflicts.

laws with α values that cluster around 2.5 also tend to fit well the distributions of violent events generated by terrorist organizations. Thus, there appear to be close parallels in the behavior of terrorist and insurgent organizations, at least with respect to the processes that generate their violent events. This empirical commonality is reassuring given the blurred distinctions between the two types of organizations (Moghadam et al., 2014) which often seem to be almost arbitrary in practice.

We now exploit the above theoretically grounded empirical findings displayed in figure 2 to make predictions about event sizes. We proceed by generating out-of-sample predictions for the expected ratios of event counts for various pairs of size ranges. We then calculate the successes and failures of these predictions. Specifically, we implement the following procedures.

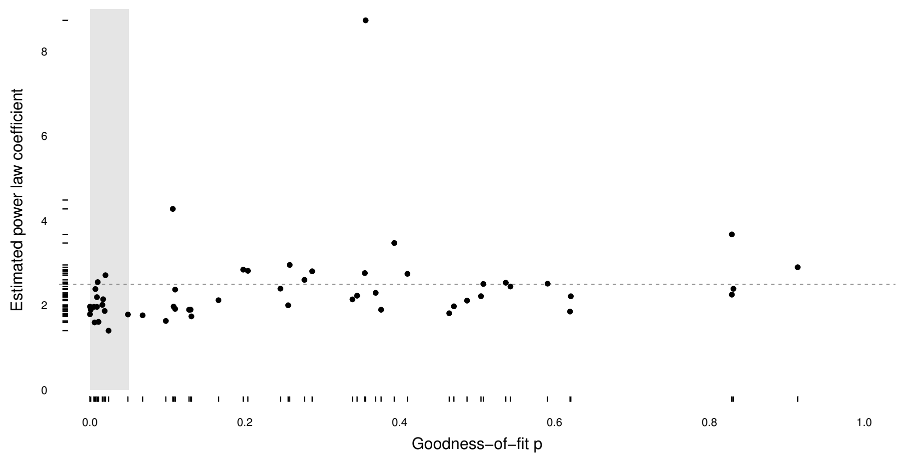

Figure 3: Estimates of α parameters versus p values for power-law hypotheses for terrorist organizations.

1. Randomly split the sample into two parts and use one third of the conflicts to generate out-of-sample predictions for the remaining two thirds of the conflicts.

2. Fit power laws to the third that were selected.

3. Order the α estimates from the selected third from smallest to largest and calculate the range of α's running from percentile 2.5 to percentile 97.5.

4. Use the lower and upper bounds of this range to predict the upper and lower bounds, respectively, of the ratios of event-size counts for various ratios of event size ranges. For example, if the lower bound for α is 2.0 then the upper bound for the ratio of the number of events of size S or greater to the number of events of size $2S$ or greater is 2 while if the upper bound for α is 3.5 then the lower bound for the same ratio of event-size ranges is about $5.7S$. The corresponding figures for S and $1.5S$ are $1.5S$ and $2.8S$ respectively.

5. Check these predictions against the data for the two thirds of conflicts that were not randomly selected. Although we could check a near-endless list of predictions we confine ourselves to just checking the ratios for which we multiply the event size by either 1.5 or 2.0.

6. Start over, taking a new draw of 1/3 of the conflicts and again testing the out-of-sample predictions on the remaining 2/3 of conflicts.

We repeat this procedure $1,000$ times. Figure 4 displays the results for this simulation exercise. For most event-size ratios the success rates exceed 60% for at least 75% of the draws of $1,000$. The best prediction performance is for the event-size ratios of $10/20$ and $20/40$ for which the median success rates are in the 80's and even the worst runs tend to score well above 60%. The worst prediction performances are when the events are either very small or very large. The relatively low success rates for small events make sense since the estimated power laws are not even meant to apply below some cut-off level s_{min}. Thus, if anything, the success rates for the low-end predictions are a bit of a bonus. Relatively weak performance at the high end also makes sense since the data on big events are sparse, providing only a thin empirical basis for prediction. Note, further, that these good prediction scores are not generally due to vacuously wide prediction intervals as the typical the intervals are around 1.2 to 3.5 and 1.4 to 8.6 for size ranges of the form S to $1.5S$ and S to $2S$ respectively (with the upper limit of 8.6, admittedly, rather high).

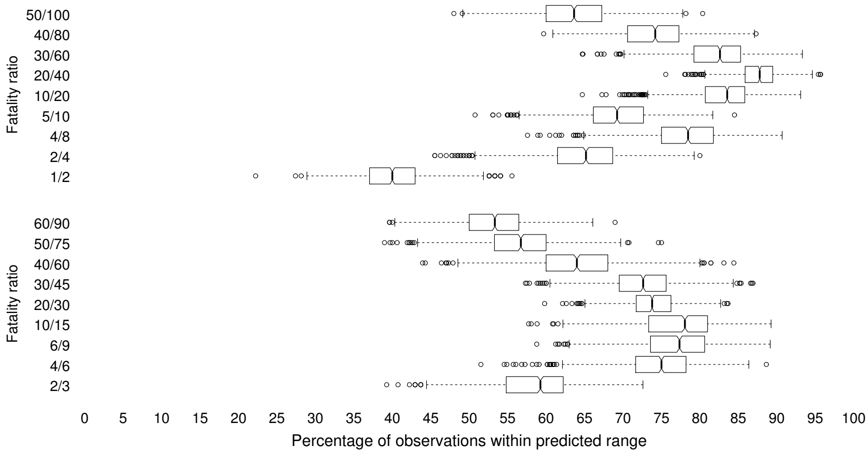

Figure 4: Boxplots for the distributions of the percentage of successful out-of-sample predictions for a variety of ratios of event-size ranges.

Figure 5 shows that out-of-sample prediction works almost as well as in-sample prediction for our power-law based scheme. The solid curves give the success rates when we use all data to generate the α range and then test the predictions (self-referentially) on the whole dataset. The grey-shaded area indicates the middle 50% of the success rates for the 1,000 out-of-sample runs.

Next we compare the performance of our power-law-based predictions with a similar scheme that uses the lognormal distribution instead. Each

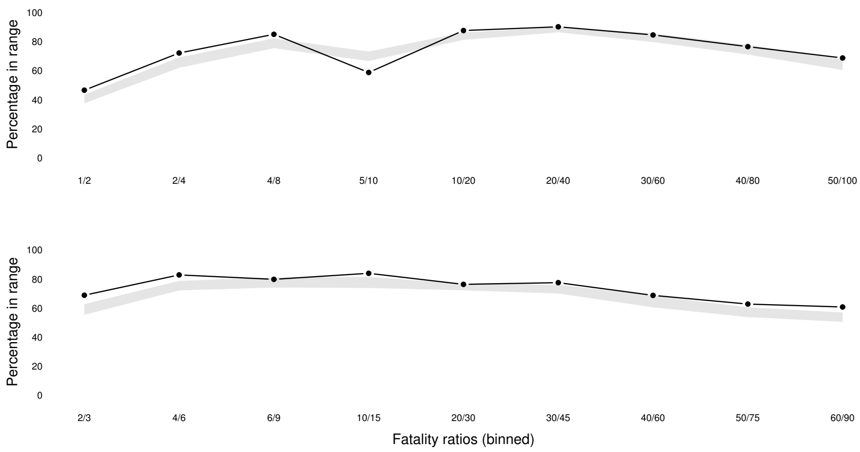

Figure 5: The success rates for in-sample predictions compared to the success rates for out-of-sample predictions. The shaded area indicates the 50% interval for the out-of-sample results.

point in figure 6 gives two statistics describing the outcome of out-of-sample predictions for a particular randomly split sample. The x-axes give the percentages of within-boundary predictions for the 10/15 ratio, ranging over all the out-of-sample conflicts. The y-axes provide a measure that combine considerations of how accurate and how unhedged, i.e., how narrow, each prediction interval is. Specifically, we define the *AH score* for a prediction interval as the inverse of the root mean squared distance from its boundary predictions (percentiles 2.5 and 97.5) to the actual 10/15 fatalities ratio. Thus, the AH score most strongly rewards prediction intervals that are both accurate, i.e., centered around the true value, and minimally hedged, i.e., narrow. Figure 5 shows the power law system beating the lognormal system: the power-law based prediction intervals have systematically higher AH scores than the lognormal-based intervals do with little or no cost to the percentage of correctly predicted ratios. The small bunch of points in the power law picture with AH scores near 0.25 are produced on simulation runs that repeatedly pick the (outlier) conflicts with very high α's.

Conclusion

We have investigated the size distribution of violent events in modern conflicts and terrorist campaigns, finding that these are generally well approximated by power laws with α coefficients clustered near 2.5, although there are exceptions. We exploit these empirical regularities in the conflict data, without ignoring

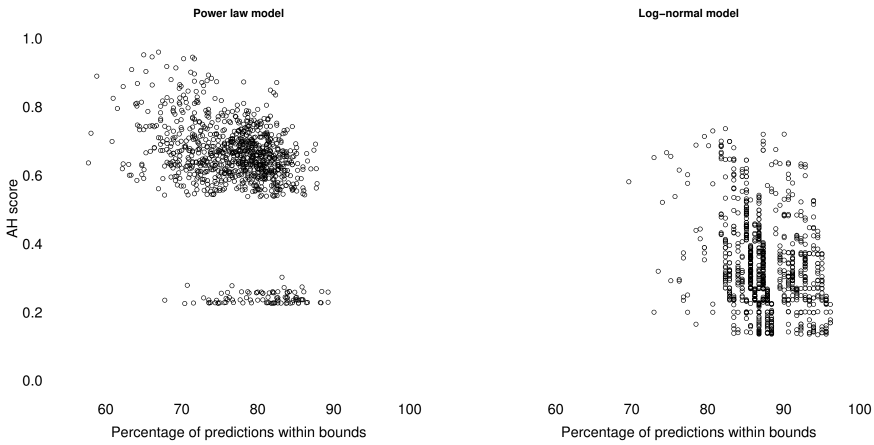

Figure 6: Prediction intervals for the power-law-based prediction system tend to be narrower than the prediction intervals based on the lognormal model at little or no cost to prediction accuracy.

the anomalies, and are able to make good predictions about the relative frequencies of violent events falling within various size classes. Our success at out-of-sample predictions indicates that our approach should work well for predicting the mixtures of event sizes in future armed conflicts. Specifically, our sample of 202 conflicts suggests a useful rule of thumb whereby the α's for future conflicts are predicted to be 2.5 with a prediction interval of 1.5 (percentile 2.5) to 4.1 (percentile 97.5). Using this insight, we can predict that for any s the ratio of the number of events of size s or more to the number of events of size $1.5s$ or more will be approximately 1.8 with a prediction interval of 1.2 to 3.5. For events of sizes s and $2s$ the analogous numbers are 2.8, 1.4 and 8.6. This level of predictability should be useful for purposes such as the planning of emergency medical care in conflict zones. More importantly, these results deepen our understanding of the fundamentals of terrorism and modern warfare and underline the potential for the coalescence-fragmentation model (Bohorquez et al., 2009; Johnson et al., 2013) to further illuminate these fundamentals. The strong parallels we find between insurgent and terrorist organizations also extend our understanding of the nature of violent conflict. These apparently different phenomena display deep common patterns that transcend their surface-level differences. Analysts of modern war and terrorism (Hammes, 2006; Robb, 2007; Moghadam et al., 2014) have been correct to broadly view these contentious situations as archetypal David versus Goliath confrontations.

References

Bohorquez, J. C., S. Gourley, A. R. Dixon, M. Spagat, and N. F. Johnson (2009, dec). Common ecology quantifies human insurgency. *Nature 462*(7275), 911–4.

Cederman, L. E. (2003). Modeling the Size of Wars: From Billiard Balls to Sandpiles. *The American Political Science Review 97*(1), 135–150.

Clauset, A. (2017). The constant hazard of a large war. *Forthcoming*.

Clauset, A., C. Shalizi, and M. Newman (2009). Power-law distributions in empirical data. *SIAM Review 51*, 661–703.

Clauset, A. and R. Woodard (2013a). Estimating the Historical and Future Probabilities of Large Terrorist Events. *Annals of Applied Statistics 7*(4), 1838–1865.

Clauset, A. and R. Woodard (2013b). Rejoind of "Estimating the Historical and Future Probabilities of Large Terrorist Events". *Annals of Applied Statistics 7*(4), 1838–1865.

Clauset, A., M. Young, and K. Skrede Gleditsch (2007). On the Frequency of Severe Terrorist Events. *Journal of Conflict Resolution 51*(1), 58–87.

Eck, K. (2012, mar). In data we trust? A comparison of UCDP GED and ACLED conflict events datasets. *Cooperation and Conflict 47*(1), 124–141.

Gillespie, C. S. (2014). Fitting heavy tailed distributions: the poweRlaw package. *Journal of Statistical Software* (July), 1–16.

Hammes, T. (2006). *The sling and the stone: on war in the 21st century*. Zenith Press.

Johnson, N., M. Zheng, Y. Vorobyeva, A. Gabriel, H. Qi, N. Velasquez, P. Manrique, D. Johnson, E. Restrepo, C. Song, and S. Wuchty (2016). New online ecology of adversarial aggregates: ISIS and beyond. *Science (New York, N.Y.) 352*(6292), 1459–1463.

Johnson, N. F., P. Medina, G. Zhao, D. S. Messinger, J. Horgan, P. Gill, J. C. Bohorquez, W. Mattson, D. Gangi, H. Qi, P. Manrique, N. Velasquez, A. Morgenstern, E. Restrepo, N. Johnson, M. Spagat, and R. Zarama (2013). Simple mathematical law benchmarks human confrontations. *Scientific reports 3*, 3463.

Moghadam, A., R. Berger, and P. Beliakova (2014). Say Terrorist, Think Insurgent: Labeling and Analyzing Contemporary Terrorist Actors. *Perspectives on Terrorism 8*(5), 2–17.

National Consortium for the Study of Terrorism and Responses to Terrorism (START) (2016). Global terrorism database [data file]. Accessed 30 June 2016.

Richardson, L. (1960). *Statistics of deadly quarrels.* Pacific Grove, California, USA: Boxwood Press.

Robb, J. (2007). *Brave new war: The next stage of terrorism and the end of globalization.* John Wiley & Sons.

Sundberg, R. and E. Melander (2013). Introducing the UCDP Georeferenced Event Dataset. *Journal of Peace Research 50*(4), 523–532.

Weidmann, N. B. (2013). The Higher the Better? The Limits of Analytical Resolution in Conflict Event Datasets. *Cooperation and Conflict 48*(4), 567–576.

Weidmann, N. B. (2015). On the Accuracy of Media-based Conflict Event Data. *Journal of Conflict Resolution 59*(6), 1129–1149.

UCD CENTRE FOR ECONOMIC RESEARCH – RECENT WORKING PAPERS

WP16/18 Stijn van Weezel: 'Short term effects of drought on communal conflict in Nigeria' December 2016
WP16/19 Sarah La Monaca and Lisa Ryan: 'Solar PV where the sun doesn't shine: Estimating the economic impacts of support schemes for residential PV with detailed net demand profiling' December 2016
WP16/20 Kevin O'Rourke: 'Independent Ireland in Comparative Perspective' December 2016
WP17/01 Roberta Cardani, Alessia Paccagnini and Stelios Bekiros: 'The Effectiveness of Forward Guidance in an Estimated DSGE Model for the Euro Area: the Role of Expectations' January 2017
WP17/02 Doireann Fitzgerald, Stefanie Haller and Yaniv Yedid-Levi: 'How Exporters Grow' January 2017
WP17/03 Igor Bagayev and Ronald B Davies: 'The Infant Industry Argument: Tariffs, NTMs and Innovation' January 2017
WP17/04 Igor Bagayev and Ronald B Davies: 'Non-homothetic Preferences, Income Distribution, and the Burden of NTMs' February 2017
WP17/05 Igor Bagayev and Ronald B Davies: 'The Impact of Protection on Observed Productivity Distributions' February 2017
WP17/06 Igor Bagayev, Ronald B Davies, Panos Hatzipanayotou, Panos Konstantinou and Marie Rau: 'Non-Tariff Barriers, Enforcement, and Revenues: The Use of Anti-Dumping as a Revenue Generating Trade Policy' March 2017
WP17/07 Simone Wegge, Tyler Anbinder and Cormac Ó Gráda: 'Immigrants and Savers: A Rich New Database on the Irish in 1850s New York' April 2017
WP17/08 Ronald B Davies and Zuzanna Studnicka: 'The Heterogeneous Impact of Brexit: Early Indications from the FTSE' May 2017
WP17/09 J Peter Neary and Cormac Ó Gráda: 'Brendan M. Walsh (1940-2016): The Economist at Work' May 2017
WP17/10 Morgan Kelly and Cormac Ó Gráda: 'Speed under Sail, 1750–1830' June 2017
WP17/11 Morgan Kelly and Cormac Ó Gráda: 'Technological Dynamism in a Stagnant Sector: Safety at Sea during the Early Industrial Revolution' June 2017
WP17/12 Kate Hynes, Yum K Kwan and Anthony Foley: 'Local linkages: The interdependence of foreign and domestic firms' June 2017
WP17/13 Cormac Ó Gráda: 'Notes on the Demography of the Famine in Ulster' June 2017
WP17/14 Sarah Parlane and Yingyi Tsai: 'Optimal Management of Supply Disruptions when Contracting with Unreliable, Risk-averse, Suppliers' June 2017
WP17/15 Orla Doyle: 'The First 2,000 Days and Child Skills: Evidence from a Randomized Experiment of Home Visiting' July 2017
WP17/16 Christopher Dixon-O'Mara and Lisa Ryan: 'Energy efficiency in the food retail sector: Barriers, drivers, and acceptable policies' July 2017
WP17/17 Andrew E Clark, Orla Doyle, and Elena Stancanelli: 'The Impact of Terrorism on Well-being: Evidence from the Boston Marathon Bombing' September 2017
WP17/18 Kate Hynes, Jie Ma and Cheng Yuan: 'Transport Infrastructure Investments and Competition for FDI' September 2017
WP17/19 Kate Hynes, Eric Evans Osei Opoku and Isabel KM Yan: 'Reaching Up and Reaching Out: The Impact of Competition on Firms' Productivity and Export Decisions' September 2017
WP17/20 Tamanna Adhikari, Michael Breen and Robert Gillanders: 'Are New States More Corrupt? Expert Opinions vs. Firms' Experiences' October 2017