

“Natural Disasters, Casualties and Power Laws: A Comparative Analysis with Armed Conflict”

Óscar Becerra¹, Neil Johnson^{2,3}, Patrick Meier^{4,5}, Jorge Restrepo^{3,6} and Michael Spagat^{7,3*}

**In alphabetical order only*

¹*Fedesarrollo, Bogotá, Colombia*

²*Department of Physics, University of Oxford, Oxford UK*

³*CERAC, Conflict Analysis Resource Center, Bogotá, Colombia*

⁴*The Fletcher School of Law and Diplomacy, Tufts University, Boston, US*

⁵*Center for Study of Civil War (CSCW), Peace Research Institute, Oslo (PRIO), Norway*

⁶*Department of Economics, Javeriana Univeristy, Bogotá, Colombia*

⁷*Department of Economics, Royal Holloway College, University of London, Egham, UK*

Abstract

Power-law relationships, relating events with magnitudes to their frequency, are common in natural disasters and violent conflict. Compared to many statistical distributions, power laws drop off more gradually, i.e. they have “fat tails”. Existing studies on natural disaster power laws are mostly confined to physical measurements, e.g., the Richter scale, and seldom cover casualty distributions. Drawing on the Center for Research on the Epidemiology of Disasters (CRED) International Disaster Database, 1980 to 2005, we find strong evidence for power laws in casualty distributions for all disasters combined, both globally and by continent except for North America and non-EU Europe. This finding is timely and gives useful guidance for disaster preparedness and response since natural catastrophes are increasing in frequency and affecting larger numbers of people. We also find that the slopes of the disaster casualty power laws are much smaller than those for modern wars and terrorism, raising an open question of how to explain the differences. We show that many standard risk quantification methods fail in the case of natural disasters.

Paper prepared for presentation at the 2006 Annual Meeting of the American Political Science Association (APSA), Philadelphia, PA, August 31-September 3, 2006.

Contact: Patrick.Meier@Tufts.Edu or M.Spagat@rhul.ac.uk.

*****PRE-CONFERENCE DRAFT ABSOLUTELY NOT FOR CIRCULATION*****

Introduction

Power law distributions are frequently found in biological, physical and social systems (Newman 2005, Buchanan 2000).¹ For example, some natural disasters such as the size versus frequency of earthquakes, forest fires and landslides follow a power law distribution. Related research suggests that casualties in whole wars (Richardson 1948, 1960; Cederman 2003), global terrorist events (Clauset and Young 2005), and in war events (Johnson et al. 2005 & 2006) follow power laws.

Natural disasters are a major threat to human security. Purvis and Busby (2004, 68) reckon that natural disasters affected roughly 188 million people per year between 1990-1999; an order of 6 times the affected population for armed conflict over the same period. Moreover, natural-disaster frequency is increasing over time (Emergency Disasters Data Base or EM-DAT).

Knowledge of casualty distributions in natural disaster events can give critical support to pre-disaster planning (Combs et al. 1999). A key component of contingency planning is the identification of resources most likely to be in demand in a post-disaster environment when disaster medical assistance teams (DMATs) and disaster mortuary operational response teams (DMORTs) are deployed.² Knowledge of the relative frequencies of disasters with various casualty counts is vital for the rational stockpiling medical supplies, assembly of rapid-response teams with the right numbers and skill mixes and insurance planning (among other things).

In this paper we analyze natural disaster casualty distributions using EM-DAT's Center for Research on the Epidemiology of Disasters (CRED) International Disaster Database, 1980-2005. We aggregate all disaster types and test for casualty power laws for individual continents and for the whole world.

¹ A key property of power laws is that their densities decline more gradually than those of many common statistical distributions (such as the Gaussian or Normal distribution): the so-called "fat tail" property. In addition, both densities and distributions of power laws follow straight lines when plotted on log-log axes. See the appendix for some basic mathematics of power laws.

² Of course, post-disaster casualty estimation is a critical component of any evolving response to a particular disaster that has already occurred (Sharma, 2001), but we do not consider this question in the present paper.

Here are our main findings. First, there is good evidence for power laws in disasters at the global level. Second, power laws are well supported by continent, except for North America and non-EU Europe. Third, the exact form (slope) of these power laws varies little from case to case, i.e., the findings are robust to continent-by-continent disaggregation. Fourth, natural disaster power laws are flatter than those arising from terrorism events on non-G7 targets (Clauset and Young, 2005) and modern war events (Johnson et al. 2005 & 2006). This means that the relative frequencies of high-casualty events to low-casualty events are higher for natural disasters than the corresponding ratios for non-G7 terrorism and modern war, i.e., natural disaster casualty distributions have fatter tails. In fact, natural disaster power laws are close to the power-law distributions that have been observed for casualties in whole wars (Richardson 1948, 1960; Cederman 2003) and for terrorism on G7 targets (Clauset and Young, 2005). Finally, because natural disaster risks follow a power law distribution, we argue that commonly used variance-based risk measures are deeply problematic in this context.

This paper is structured as follows: We first review the literature on casualty distributions in natural disasters to provide context for the analysis that follows. Next we describe the data we use and the methodology we apply in our study. We then present our findings, relate them to the literature on war casualties, armed conflict and terrorism, and critique the standard risk quantification approaches within the context of our finding. Finally, we consider the policy implications for disaster risk management and outline possibilities for future research.

Literature Review: Casualty Distributions and Natural Disasters

Blong and Radford (1993), in a pioneering study, studied a record of disasters in the Solomon Islands to assess the likely mixture of future events and contribute to the development of risk and mitigation strategies. They used data developed by the Australian Development Assistance Bureau on 209 disaster events, 27 of them occurring before 1900, including cyclones, droughts, earthquakes, floods, storms, landslides, tsunamis and volcanic eruptions. The authors' plot of the cumulative distribution function for deaths in their

natural-hazard events on a log-log scale very much resembles a straight line. So the casualty data for the Solomon Islands over a long period is suggestive of a power law, although the authors do not mention or test for one.

Guzzetti (2000) studies a wide array of landslide data (not from CRED) going back very far in time. He covers the periods 1410-1999 and 1950-1999 for Italy and similar but varying periods for Canada, the Alps, Japan, Hong Kong and China.³ Through the inspection of log-log plots Guzzetti (2000) finds that the cumulative frequency of landslide events plotted against their casualty consequences “can be approximated by power laws”.

Jonkman (2005) uses the CRED dataset to study to study the distribution of killings in global events, 1975-2001, disaggregating by disaster type including floods, droughts, earthquakes, famines, windstorms and epidemics but not by continent as we do. For each disaster type he plots the global frequency of events with N or more people killed against N on a log-log scale. Only the curve for earthquake casualties resembles a power law and Jonkman does not pursue this question. The near-absence of straight lines in these plots, together with our own straight-line results for all disasters combined, raises a puzzle as to why power laws arise globally and by continent for all disaster types but not for individual disaster types.

To summarize, we use completely different data compared to Blong and Radford (1993) and Guzzetti (2000). We use the same dataset as Jonkman (2005) but disaggregate by continent rather than by disaster type. We find quite robust power-laws evidence whereas Jonkman does not. We also contribute by introducing formal statistical testing for casualty power laws into the disaster literature. We relate our finding to the literature on war casualties, armed conflict and terrorism and we discuss the insurance implications of our finding.

Data and Methodology

³ There are, of course, significant quality issues arising in the use of such deep historical data. Nevertheless, study is highly enterprising and interesting.

Our natural-disaster data is derived from the EM-DAT International Disaster Database maintained by the Center for Research on Epidemiology of Disasters (CRED) in cooperation with the United States Office for Foreign Disaster Assistance (OFDA). The database defines a disaster as meeting at least one of the following criteria: 10 or more deaths, 2000 or more affected people for droughts and famine or 100 or more for other disasters, a government disaster declaration, or a plea for international assistance. We analyze all disasters combined, 1980-2005, both globally and disaggregated by continent.

Disaster data is necessarily imperfect and the quality of the CRED data is known to vary across years and disaster types. However, according to a quality assessment study by Guha-Sapir and Below (2002), recent CRED data compares well in quality to proprietary datasets and is the most reliable and comprehensive global dataset in the public domain. Dilley et al. (2005) hesitate to draw on CRED data to identify total mortality risk, but are quite comfortable using CRED to analyze relative risk. This is encouraging for our work since power law analysis is specifically about estimating relative magnitudes since the latter also focuses on relative magnitudes.

The ideas behind our estimation and testing procedures are as follows (details are in the appendix). We estimate two parameters, x_{min} and α , the minimum casualty level above which the power law is supposed to hold, and the exponent of this estimated power law respectively. We obtain these through an iterative procedure that starts with $x_{min} = 1$, estimates the corresponding maximum-likelihood α and then repeats the procedures for $x_{min} = 2, 3, \dots$, settling on the (x_{min}, α) pair that gives the best estimated power-law fit to the data according to the Kolmogorov-Smirnov statistic. We then test two hypotheses using Monte Carlo methods to generate the test statistics: that the data is generated from the estimated power-law curve and that the data is generated from the best-fitting lognormal distribution. The lognormal is a commonly used fat-tailed distribution and, therefore, a natural comparator. When we fail to reject the power-law hypothesis but do reject the lognormal hypothesis we consider the evidence for a power law to be strong.

Results and Discussion

We summarize our findings for the whole world and by continent in Table 1. For each case we give the number of disasters in the sample, the maximum number of casualties for these events, our x_{min} and α estimates, the number of observations greater than or equal to x_{min} , and the p-values for the Kolmogorov-Smirnov tests of the power law (PL) and lognormal (LN) hypotheses. In addition, Table 1 gives population, population density (per square kilometer), and average GDP per capita adjusted for purchasing power parity.⁴

Casualties in	Obs	xmax	\alpha	xmin	obs>xmin	KS (Ho: PL)	KS (Ho: LN)	Population	Pop density	GDP per cap
World	7476	1800090	1.7304	177	995	0.434	<0.0001	6,525,486,603	49.81	11,479.20
Americas	1808	1800090	1.7451	48	355	0.878	0.0180	893,496,906	22.96	12,198.00
North America	596	7060	2.1342	77	66	0.857	0.3408	331,672,307	17.71	34,540.00
Center/South America	1212	1800090	1.6898	35	329	0.857	0.0100	561,824,599	27.83	9,715.56
North America (with México)	730	39704	2.0877	86	89	0.932	0.3304	439,121,832	21.26	30,450.00
C/South America (without México)	1078	1800090	1.6647	29	313	0.673	0.0084	454,375,074	24.88	9,709.09
Asia and Oceania	3135	277715	1.6934	215	527	0.745	<0.0001	3,920,701,817	101.37	9,190.14
Africa	1519	300000	1.6631	39	432	0.163	<0.0001	910,849,725	30.56	3,916.07
Europe	1011	61080	1.599	9	387	0.740	0.0008	800,438,155	33.91	22,706.12
EU countries	454	20000	1.7258	9	171	0.611	0.0008	454,809,846	118.73	26,173.91
non EU countries	557	61080	1.7375	68	85	0.976	0.1664	345,628,309	17.48	19,638.46

Table 1. Estimation and testing results globally and by continent.

In most cases we find good evidence for power laws with α between 1.6 and 1.75. In these other cases rejection of the lognormal is always with confidence above 98% and almost always well over 99%. Africa is the only continent where we come anywhere close to rejecting a power law at a standard significance level. So there is a strong robustness to the findings.

The two exceptions are non-EU Europe where we cannot clearly reject the lognormal distribution and North America where we are even further from being able to reject log-normality the estimated α for the power law exceeds 2.1. One possible explanation for these deviations is simply that these two cases have the smallest sample sizes. A second possibility is that the different results are related to the fact that these two cases also have the lowest population densities. It appears that we can rule out explanations based on per capita income; although non-EU Europe and North America are relatively very rich so is EU

⁴ We took demographic indicators from the International Data Base of the U.S. Census Bureau (<http://www.census.gov/ipc/www/idbnew.html>) and GDP per capita from the CIA World Factbook (<http://www.cia.gov/cia/publications/factbook/>). The GPD per capita values for the majority of countries are for 2005 but for a small group of countries there are estimates for earlier years.

Europe which behaves much like the poor countries in its natural disaster casualty distribution.⁵

Figures 1 to 2 illustrate the power law estimation results. Figure 1a depicts our global result while Figures 1b through 1e represent distributions by individual continent. Figures 2a through 2f depict results by specific region. For each region, the empirical points are blue dots and the estimated line is green and dotted. We mark the estimated x_{min} with a vertical line that is red and dotted and we display the estimated numbers for x_{min} and α . The Figures confirm the close fit of the estimates.

These power-law findings for casualties in natural disasters fit in nicely with established results for whole wars, terrorist events and events in individual modern wars. Richardson (1948, 1960) and Cederman (2003), treating an entire war similarly to the way we treat a natural disaster in this paper find power laws with alphas, slightly higher but very close to our estimated α 's. Clauset and Young (2005) present similar results using terrorist attacks on G7 targets as the basic unit of analysis. In contrast, Clauset and Young (2005) and (Johnson et al. 2005 & 2006) find power laws with much higher α 's, around 2.5, for terrorist attacks on G7 targets and events in modern wars respectively. Some of these papers have proposed models to explain their results. It is an open question whether a model can be built to illuminate our natural disaster results.

⁵ We did not notice any other useful patterns in the relationships between the economic/demographic indicators and our size distribution results.

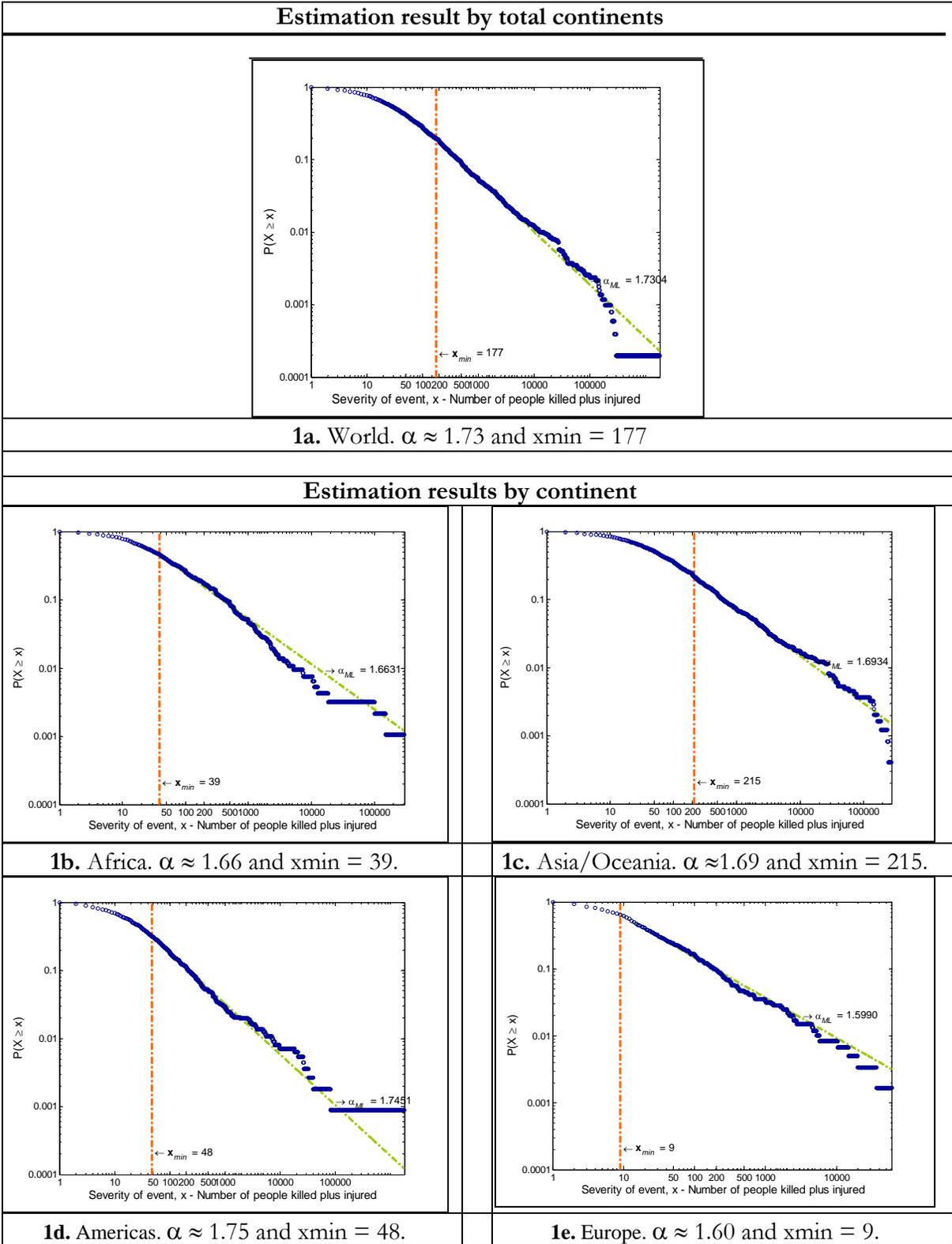


Figure 1: Estimation results. These graphs depict the casualty distribution across all disasters for all countries in the World, and disaggregated by continent.

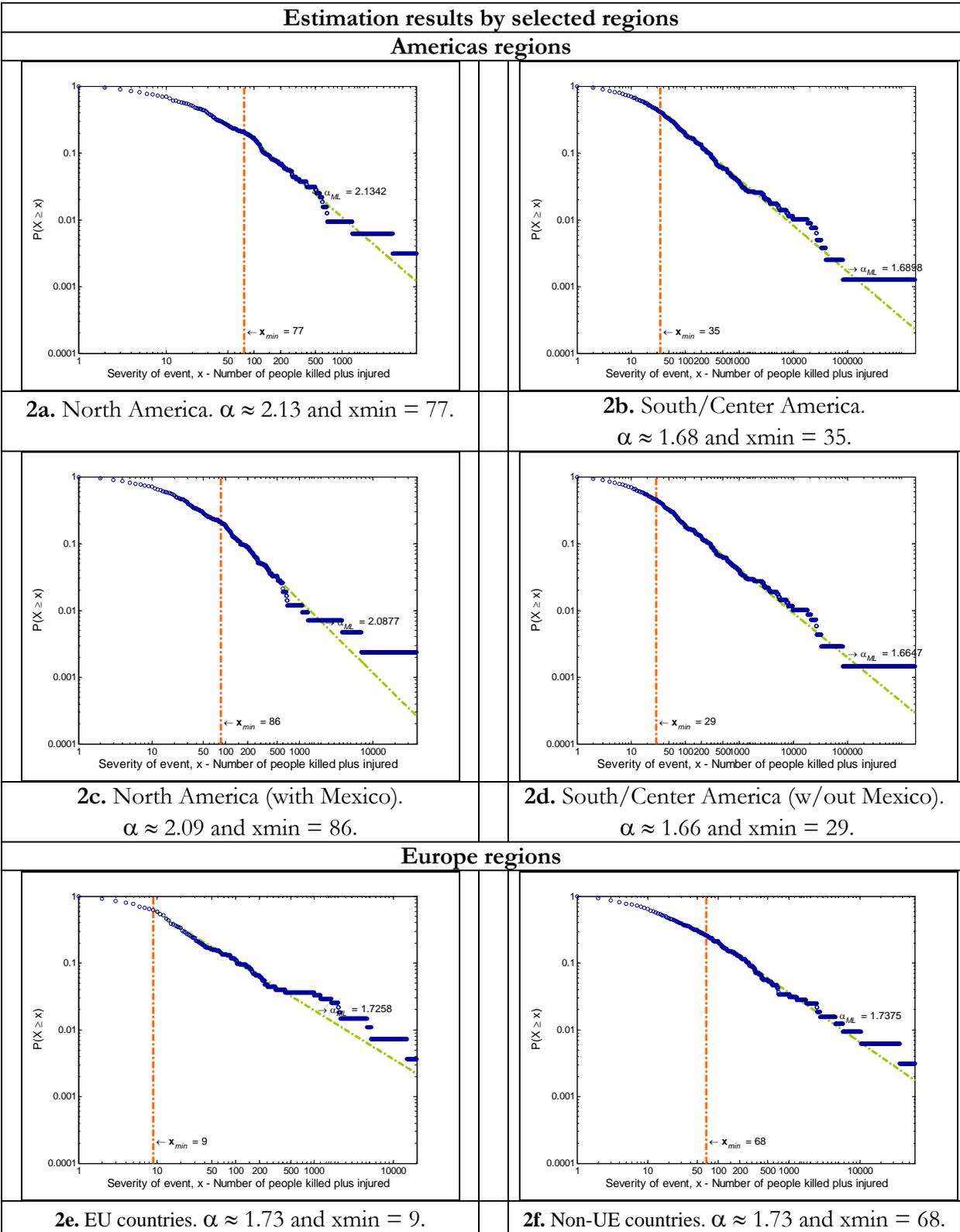


Figure 2: Estimation results. These graphs depict the casualty distribution across all disasters for all countries disaggregated by specific regions.

Our findings have important implications for standard approaches to risk evaluation. Jonkman et al. (2003) give a broad survey of quantitative methods for assessing and limiting risks. There are a very wide range of risk measures in use. Generally, these are based either on some combination of the mean and variance of the event distribution or they focus in some way on the tail of the distribution. However, power law distributions have infinite variance whenever α is below three and all our estimated α 's almost always in the range of 1.6 to 1.75, i.e., well below 3. It is true that casualty distributions do not, in reality, extend all the way to infinity; there are absolute limits to the number of people who can be killed in a natural disaster event. Nevertheless, the infinite variance problem will manifest itself empirically in high sensitivity of empirical variance measures to presence or absence of a small number of big events. This problem is intuitive. Power laws have fat tails so high-casualty events are fairly common. But the exact proportion of high-casualty events present in a dataset at a given point in time will fluctuate rather erratically causing high variance in any empirical estimate of the true variance of the underlying distribution. Thus, variance-based risk measures are highly unstable.

A second standard approach to risk assessment base on tail behavior is the so-called FN criterion. In the present context this would mean verifying that the probability of having x casualties or more is always less than $\frac{C}{x^2}$ where C is a fixed constant. However, if x follows a power law distribution with exponent α then this criterion reduces to $\frac{1}{x^{\alpha-1}} < \frac{C}{x^2}$ for all x which is impossible when α is below 3 as it always is in our estimates.⁶ Again, the problem here is intuitive. The criterion requires that event probabilities decline with their severity at a rate that is inconsistent with the fat-tail property of a power law.

⁶ More concretely when $\alpha = 1.7$, as in our natural disaster results, the inequality becomes $\frac{1}{x^{.7}} < \frac{C}{x^2}$ which of course will not hold for sufficiently high-casualty events.

Policy and Further Research

Our work suggests that there are strong regularities in casualty patterns in the broad aggregate character of natural disasters as a whole. This pattern holds up quite well on most continents. Moreover, the specific pattern is the fat-tailed power law. This means that we should expect a steady stream of large-casualty natural disaster events. Large scale disasters are not anomalies. They fall well within established patterns. Disaster risk-management planning must take heed of this fact, which justifies investments being made to develop more reliable disaster early warning systems for effective and early response.

The emergence of a power law distribution within an event list that mixes together many disaster types suggests, to a significant degree, that all disasters with x casualties will resemble one another and disaster response planning can proceed accordingly. Of course, this casualty commonality can only be pushed so far. At some stage there must be more specific planning for more specific types of disaster.

On the theoretical level our findings raise a question about whether aggregate patterns for all disasters combined can be explained within a single framework. More importantly, can such a theory provide concrete insights and recommendations to improve disaster management and response?

Finally, standard methods for quantifying risks fail badly in the natural disaster context. It is urgent to rethink these approaches in order to rationalize actual preparations.

Appendix

The probability density function for a power law in our context:

$$(1) \quad p(x) = Cx^{-\alpha}$$

where x is the number of casualties in a given event and C is a constant such that the probability of all possible outcomes (x) sums to 1.⁷ In equation (1), the probability that an event results in a number of casualties equal to x is decreasing in x with the decline depending on a “tail parameter” α where α is inversely related to the fatness of the tail. By taking the logarithm of both sides, equation (1) is expressed as:

$$(2) \quad \log p(x) = c - \alpha \log x$$

That is, if the distribution of casualties belongs to the power-law class, $\log x$ will be linear in $\log p(x)$. So a log-log graph $p(x)$ against x is a straight line as is a log-log graph of $P(X > x)$ (1 minus the cumulative distribution function) against x .

Certain considerations must be taken into account when estimating the α coefficient. First, the number of casualties is a variable of count data; it can take only integer values. Second, empirical work shows that it is uncommon for the entire distribution to follow a power law. Indeed, several researchers have found that only the ‘tail’ of the distribution, i.e. the right-part of the distribution over a given value which we call x_{min} , follows a power law.⁸ In consequence, we use the power law distribution for discrete data above x_{min} , as done by Clauset and Young (2005). This distribution takes the form:

⁷ Formally, it means that C is a constant such that $\int_1^{\infty} Cx^{-\alpha} dx = 1$.

⁸ See, for example, Price (1965), Shiode and Batty (2000), Clauset and Young (2005) and Newman (2005) among others.

$$(3) \quad p(x) = \frac{x^{-\alpha}}{\zeta(\alpha, x_{min})} \quad x \geq x_{min}$$

where $\zeta(\alpha, x_{min}) = \sum_{k=x_{min}}^{\infty} k^{-\alpha}$ is the incomplete Riemann zeta function. Thus, our estimation procedure identifies the optimal values of two parameters, x_{min} and α . We follow the procedure suggested by Clauset and Young (2005), which estimates the α parameter using maximum likelihood procedures (Johnson and Kotz 1969, 240), and identifies x_{min} through minimization of the Kolmogorov-Smirnov (KS) goodness of fit test. This tests the null hypothesis that the data belong to a specific distribution by comparing the empirical (S_x) and theoretical (F_x) distributions:

$$(4) \quad D = \max |F_x - S_x|$$

While the optimal x_{min} guarantees that this distance is the minimum, it does not necessarily imply statistical significance. As a result, we implement a Monte Carlo approach to obtain the p-value of this test. After obtaining the optimal x_{min} and α values, we use the KS test the hypothesis that the data follow a power law distribution, as well as the hypothesis that the data follow the lognormal distribution, another popular right-skewed and fat-tailed distribution.⁹

⁹ For further details above estimation procedure, see Johnson *et al.* (2005).

References

Blong R.J., and D.A. Radford, "Deaths in Natural Hazards in the Solomon Islands," published in *Disasters* (1993) 17: 1-11.

Buchanan M., *Ubiquity: The Science of History*, (Weidenfeld & Nicholson, 2000).

Burton I., Kates R.W., and White G.F., *The Environment as a Hazard, Second Edition*, (New York/London: Guilford Press, 1993).

Cardona O., "The Need for Rethinking the Concepts of Vulnerability and Risk from a Holistic Perspective: A Necessary Review and Criticism for Effective Risk Management," Bankoff, G., Frerks, G. and D. Hilhorst. (2004). *Mapping Vulnerability: Disasters, Development & People* (Earthscan, London).

Cederman L., "Modeling the Size of Wars: From Billiard balls to Sandpiles," *Amer. Pol. Sci. rev.*, 97, 135-50 (2003).

Clauset A. and Young M. "Scale invariance in Global Terrorism", e-print physics/0502014 at <http://xxx.lanl.gov> (2005)

Combs D.L., L.E. Quenemoen, R.G. Parrish, and J.H. Davis, "Assessing Disaster-attributed Mortality: Development and Application of a Definition and Classification Matrix." *International Journal of Epidemiology*, vol. 28, pp. 1124-1129 (1999).

Dilley M., Chen R., Deichman U., Lerner-Lam A., Arnold M. with J. Agwe, Buys, P., Kjekstad, B. L. and G. Yetman. (2005). "Natural Disaster Hotspots: A Global Risk Analysis," *Disaster Risk Management Series* (The World Bank Hazard Management Unit, Washington DC).

Espinosa-Aranda, J. M, Jimenez, A., Ibarrola, G., Alcantar, F., Aguilar A., Inostroza M., Maldonado S., and R. Higareda, "The Seismic Alert Systems in Mexico City and the School Prevention Program," in Zschau, J. and A. Kueppers eds. (2002). *Early Warning Systems for Natural Disaster Reduction* (Springer Verlag, Berlin).

Gilbert C., "Studying Disaster: Changes in the Main Conceptual Tools," in Quarantelli, E. L. ed. (1998). *What is a Disaster? Perspectives on the Question* (Routledge London and New York).

Guhar-Sapir, D. and R. Below. "Quality and accuracy of disaster data: A comparative analyse of 3 global data sets." Working paper prepared for the Disaster Management facility, World Bank, Brussels CRED, 2002.

Guzzetti, F. "Landslide Fatalities and the evaluation of landslide risk in Italy" *Engineering Geology* (2000) 58: 89-107

Heijmans A., "From Vulnerability to Empowerment," in Bankoff, G., Frerks, G. and D. Hilhorst. (2004). *Mapping Vulnerability: Disasters, Development & People* (Earthscan, London).

Johnson N., Spagat M., Restrepo J., Bohorquez J., Suarez N., Restrepo E., and Zarama R., "From Old Wars to New Wars and Global Terrorism," (2005) published on-line at: <http://xxx.lanl.gov/abs/physics/0506213>

Johnson N., Spagat M., Restrepo J., Becerra Ó., Bohorquez J., Suarez N., Restrepo E., and Zarama R., "Universal patterns underlying ongoing wars and terrorism," (2006) published on-line at: <http://xxx.lanl.gov/abs/physics/0605035>.

N.L. Johnson, and S. Kotz., *Univariate discrete distributions*. John Wiley & Sons, New York. (1992)

Jonkman, S.N. "Global Perspectives on Loss of Human Life Caused by Floods," *Natural Hazards* (2005) 34: 151-175;

Kelman I., "Disaster Deaths," access on July 8, 2006, available at: <http://www.ilankelman.org/disasterdeaths.html>

Kirigia, J.M., L.G. Sambo, W. Aldis, and G.M. Mwabu. 2004. "Impact of disaster-related mortality on gross domestic product in the WHO African Region". *BMC Emergency Medicine*, vol. 4, no. 1, available at: <http://www.biomedcentral.com/1471-227X/4/1>

Newman M.E.J., "Power Laws, Pareto distribution and Zipf's law," *Contemp. Phys.* 46, N.5, 323-351 (2005).

Pulwarty R., Broad K., and T. Finan, "El Nino Events, Forecasts and Decision-Making," in Bankoff, G., Frerks, G. and D. Hilhorst. (2004). *Mapping Vulnerability: Disasters, Development & People* (Earthscan, London).

Price D. J. "Networks of Scientific Papers", *Science* 149, 510-15 (1965).

Purvis, N. and Busby J., "The Security Implications of Climate Change for the UN System," *Environmental Change and Security Project*, May 2004.

Richardson L.F., "Variation of the Frequency of Fatal Quarrels with Magnitude, *Amer. Stat. Assoc.* 43, 523-46 (1948).

Richardson L.F., *Statistics of Deadly Quarrels*, eds. Q. Wright and C.C. Lienau (Boxwood Press, Pittsburgh, 1960).

Sharma, V. "Gujarat Earthquake: Some Emerging Issues", *Asian Disaster Management* Vol. 7, No. 2 & 3, April-September 2001.

Shiode N. and Batty M. "Distributions in Real and Virtual Worlds" Working paper series of the Centre for Advanced Spatial Analysis. 19 (2000)

The Earth Institute, *Risk Analysis Reports Over Half of World's Population Exposed to One or More Major Natural Hazards*. The Earth Institute, Columbia University, 2005.