LETTERS

Common ecology quantifies human insurgency

Juan Camilo Bohorquez[1], Sean Gourley[2], Alexander R. Dixon[3], Michael Spagat[4] & Neil F. Johnson[2]

Many collective human activities, including violence, have been shown to exhibit universal patterns[1–19]. The size distributions of casualties both in whole wars from 1816 to 1980 and terrorist attacks have separately been shown to follow approximate power-law distributions[6,7,9,10]. However, the possibility of universal patterns ranging across wars in the size distribution or timing of within-conflict events has barely been explored. Here we show that the sizes and timing of violent events within different insurgent conflicts exhibit remarkable similarities. We propose a unified model of human insurgency that reproduces these commonalities, and explains conflict-specific variations quantitatively in terms of underlying rules of engagement. Our model treats each insurgent population as an ecology of dynamically evolving, self-organized groups following common decision-making processes. Our model is consistent with several recent hypotheses about modern insurgency[18–20], is robust to many generalizations[21], and establishes a quantitative connection between human insurgency, global terrorism[10] and ecology[13–17,22,23]. Its similarity to financial market models[24–26] provides a surprising link between violent and non-violent forms of human behaviour.

The political scientist Spirling[27] and others[9,10] have correctly warned that finding common statistical distributions (for example, power laws) in sociological data is not the same as understanding their origin. Possible political, ideological, cultural, historical and geographical influences make conflict arguably one the 'messiest' of all human activities to analyse. Mindful of these challenges, yet inspired by recent studies of human dynamics[1–11,17,28], we analyse the size and timing of 54,679 violent events reported within nine diverse insurgent conflicts, placing equal emphasis on both finding and modelling common patterns. Such insurgencies typify the future wars and threats faced by society[18,19].

Our data sources are real-time media databases, official (government and non-governmental organization) reports and academic studies. Supplementary Information provides details, plus data-set extracts. The event data from different conflicts were compiled by different researchers, often with cross-checking by independent research teams, thereby reducing systematic collection or recording biases. Comparison of event accounts across a wide range of sources[12,29] reduces potential media bias and mistaken aggregation (for example, misreporting two events of sizes x_1 and x_2 as one event of size $x_3 = x_1 + x_2$), which would create significant errors in a tail-dependent estimate such as a power-law slope. We focus on measuring deaths, because injuries are harder to cross-check. However, where possible, we check that our conclusions are robust to the inclusion of injuries.

Figures 1 and 2 show our empirical findings for event size, whereas Fig. 3 shows event timings. Our model (described later and shown schematically in Fig. 4) provides a quantitative explanation of these findings by treating the insurgent population as an ecology of dynamically evolving, decision-making groups, in line with several recent sociological hypotheses[18–20]. In addition to explaining the ubiquity of approximate power laws in the event size distribution and the apparent central role of the 2.5 exponent value (Fig. 1), it explains the conflict-dependent deviations beyond a power law (see green curves in Fig. 2). Furthermore, the same model framework also explains the common burstiness in the distribution of event timings that we observe across insurgent conflicts (see black curves in Fig. 3).

Figure 1 gives exponent values, obtained by applying Clauset *et al.*'s[9,10] established methodology for estimating discrete power-law distributions, $p(x) \approx x^{-\alpha}$, for $x \geq x_{\min}$ where $x_{\min}$ is estimated together with α. In all cases we cannot reject the hypothesis that the size distribution of the events follows a power law, but we can reject log-normality. Four detailed examples are shown in Fig. 2. Following our preliminary 2005 results for Iraq and Colombia, we had suggested[12] that other insurgent wars might be clustered around $\alpha = 2.5$. All the insurgent wars that we have analysed support this hypothesis. By contrast, we find that the Spanish Civil War and the American Civil War—neither of which are considered insurgent—each give distributions where log-normal can not be rejected, and where even the best-fit α value is much smaller (near 1.7, which is the value for the aggregated sizes of conventional wars[9]). This finding provides quantitative support for claims circulating in social science[18,19]

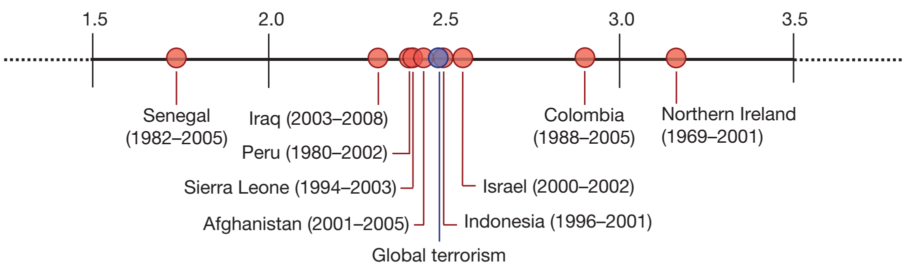

Figure 1 | Power-law exponents. Value α for power law $p(x) \approx x^{-\alpha}$ deduced from the empirical distributions of event size x (that is, the number of casualties) for insurgent conflicts. Statistical procedures follow refs 9 and 10. Blue dot shows the value 2.48 for distribution of total size of global terrorist events, from Clauset *et al.*[10]. The years in parentheses describe the empirical data set range used to deduce α, not the actual conflict duration.

[1]Department of Industrial Engineering and CEIBA Complex Systems Research Center, Universidad de Los Andes, Bogota, Colombia. [2]Complex Systems Group, Physics Department, University of Miami, Florida 33126, USA. [3]Cavendish Laboratory, Cambridge University, Cambridge CB3 0HE, UK. [4]Department of Economics, Royal Holloway College, University of London, Egham, Surrey TW20 0EX, UK.

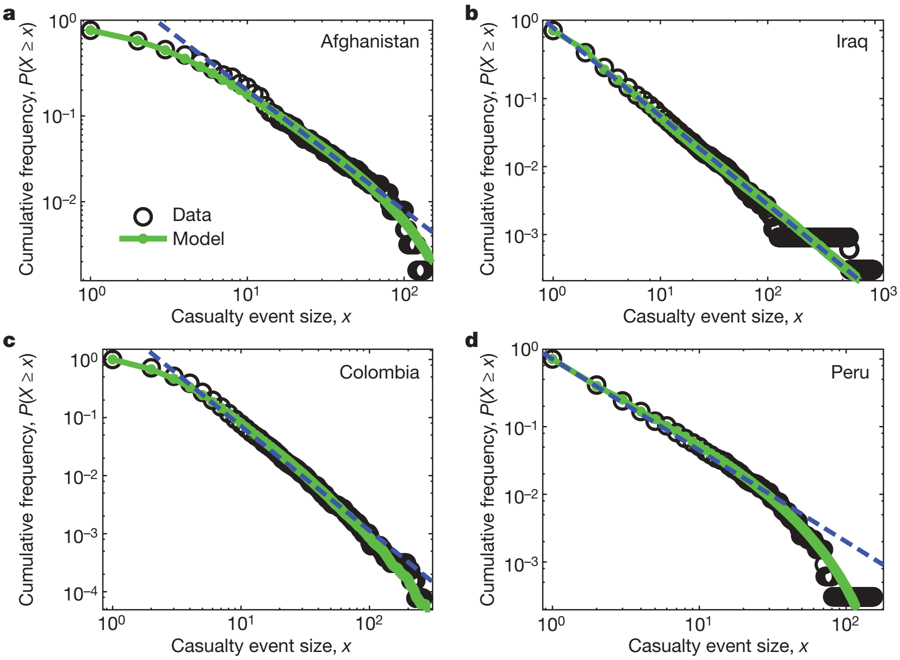

Figure 2 | Size of events. **a–d**, Log–log plot of the complementary cumulative distribution of event size $P(X \geq x)$ (that is, the probability of an event of size greater than or equal to x) for four conflicts from Fig. 1. Horizontal axis shows event size x, namely the number of casualties. Solid green curves show the results from our model. Blue dashed line is a straight line guide to the eye, not a power-law fit.

that insurgent wars represent 'open-source'[18], 'fourth-generation'[19] warfare, with qualitatively different dynamics from traditional wars. Several trivial explanations of the data can be ruled out, such as proportionality to city size[10].

Figure 3 demonstrates a common burstiness in the distribution for the number of events per day, n, irrespective of size. As explained in the Methods and Supplementary Information, we compare the distributions over daily event counts for different epochs within the four modern conflicts for which we have such data, against control distributions ('random war') obtained by randomizing event occurrences within each epoch. The data for each conflict (green circles) deviate from its random war (dashed curve) in a similar way: the real war exhibits an overabundance of light days (that is, days with few attacks) and of heavy days (that is, days with many attacks), but a 'lack' of medium days compared with the random war (see lower panel). By considering subsets of days, we have determined that these features are not just an artefact of a variation in attack volume across days of the week (for example, Fridays; see Supplementary Information). Interestingly, this

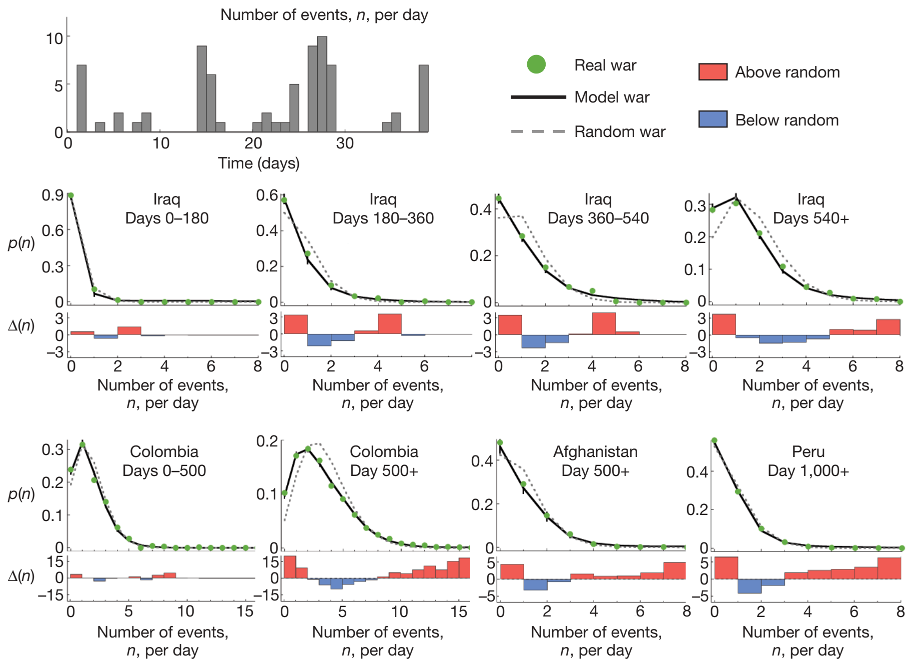

Figure 3 | Timing of events. A time series with $n = 0, 1, 2, \ldots$ events per day. Green circles show distribution $p(n)$ for the number of days with n events in actual conflict. Dashed lines represent average values for random wars. Solid lines denote average distributions calculated from 10,000 realizations of our model (Fig. 4). Histograms below represent differences $\Delta(n)$ between real and random wars, in units of standard deviations from the mean. Error bars for random wars, namely one standard deviation from the mean of 10,000 shufflings, are shown but are small. Error bars for model wars demonstrate a small spread in run outcomes.

burstiness has become more pronounced over time for the wars in both Iraq and Colombia, suggesting that they have become less random as they have evolved. These findings are insensitive to the precise specification of the epochs within a given conflict.

Our model framework in Fig. 4 incorporates two key features: (1) ongoing group dynamics within the insurgent population (for example, as a result of internal interactions and/or the presence of an opposing entity such as a state army); (2) group decision-making about when to attack based on competition for media attention. Within this framework, we find that mechanism (1) dictates the features observed in Figs 1 and 2 whereas mechanism (2) dictates those in Fig. 3. Mechanism (1) is consistent with recent work on human group dynamics in everyday environments[28], and with current views of modern insurgencies as fragmented, transient and evolving[18]. Mechanism (2) is consistent with comments by former US Senior Counterinsurgency Adviser David Kilcullen, who noted[20] that when insurgents ambush an American convoy in Iraq, '... they're not doing that because they want to reduce the number of Humvees we have in Iraq by one. They're doing it because they want spectacular media footage of a burning Humvee.' We consider the insurgent population as having an overall strength N comprising human combatants, information, resources and weapons—though, for simplicity, one can think of N humans. N is continually being re-partitioned through coalescence and fragmentation processes, thereby producing an ecology of groups. A group's strength at time-step t determines the number of human casualties x it would produce if it decided to engage in an event at that time-step. We take N to be approximately constant over time, though our main conclusions are unchanged if N evolves slowly with small fluctuations.

These two coexisting dynamic mechanisms generate rich time series that can explain the numbers of events of different sizes at each time-step. However, because the data in Figs 1 and 2 are time-aggregated whereas those in Fig. 3 are size-aggregated, we can provide far more insightful explanations by using simplified versions that treat the respective non-dominant mechanism in an averaged way. Consider first the simple situation in which the group coalescence and fragmentation processes in the insurgent population are represented by probabilities[18]. The fragmentation probability v_{frag} is taken to be small ($\sim$1%) to mimic the infrequent situation in which a group member suddenly senses imminent danger and the entire group scatters. If fragmentation does not occur, the group may coalesce with another group with probability v_{coal}. This mimics the situation in which two individuals initiate a communications link between them of arbitrary range (for example, a mobile phone call), and hence their respective groups of strength s_1 and s_2 act in a coordinated way with strength $(s_1 + s_2)$. Because these two processes can be triggered by any particular constituent group member at any time, the probability that it affects a specific group should be proportional to s (refs 24–26). Treating mechanism (2) in an averaged way, we assume that all groups are equally likely to be involved in an event over time. This is consistent with the time-averaged behaviour of the full decision-making model (see later). The time-averaged distribution of group strengths s therefore acts like the distribution of event sizes x (ref. 12), resulting in a steady-state approximate power-law distribution whose analytic solution $\alpha = 2.5$ (refs 25, 26) is within the empirical bounds of Clauset *et al.*'s total value of 2.48 ± 0.07 for global terrorism[10]. This analytically obtained theoretical value[25,26] of 2.5 is robust to many model generalizations[21,25,26] (for example, coalescence of multiple groups, fragmentation into groups larger than one), thereby offering an explanation for the observed bunching of the empirical values around 2.5 in Fig. 1.

Invoking a more realistic mechanism for grouping dynamics than simple probabilities (see Supplementary Information), we find that our model framework can explain not only the approximate power-law behaviour and central role of the 2.5 exponent (Fig. 1) but also the behaviours beyond power law observed in Fig. 2. Accounting explicitly for an opposing population (for example, state army) with total strength N_{B}, the coalescence and fragmentation are now caused by the interactions between groups. The casualties produced by clashes between opposing groups can then be used to obtain the event size distributions (green curves in Fig. 2). Full details are given in the Methods and Supplementary Information, with the four model parameters for each conflict (total insurgent strength N_{A}, total state strength N_{B} and casualty scales C_{S} and C_{L}). When two opposing groups meet they fight, with some members of both groups killed and the smaller group fragmenting. C_{S} (C_{L}) sets the scale for the

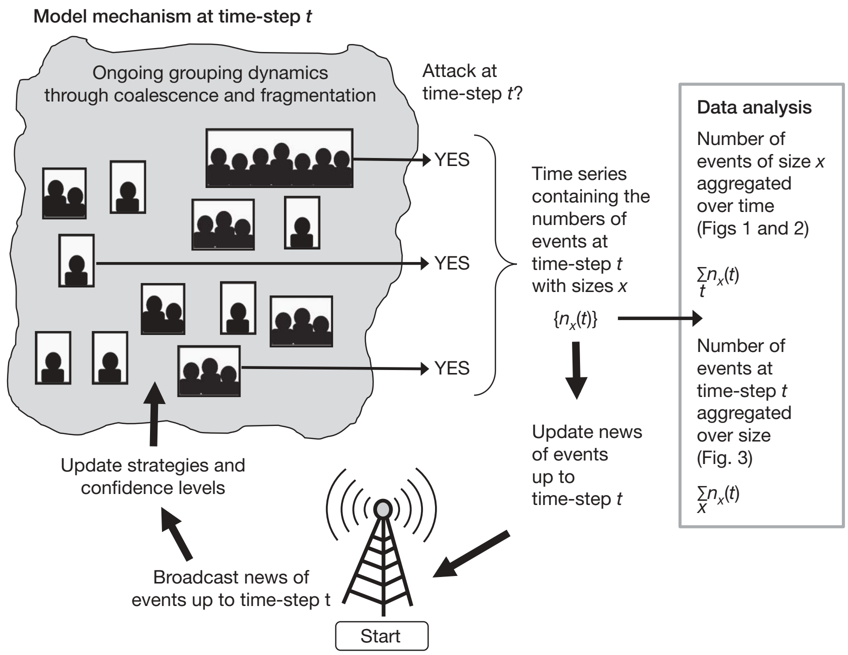

Figure 4 | Model framework for insurgency. The insurgent population comprises an overall strength N, distributed into groups with diverse strengths at each time-step t. This distribution changes over time as groups join and break up. Dark shadows indicate strength, and hence casualties that can be inflicted in an event involving that group. Figures 1 and 2 are derived from the number of events of size x aggregated over time. Figure 3 is derived from the number of events at a given time-step aggregated over size.

number of the smaller (larger) group's members destroyed. As C_S is increased, the model deviates increasingly from a straight line at low x, suggesting that Afghanistan and Colombia share the following similarity: in a clash in which the insurgent group is the smaller group, this insurgent group takes heavier relative losses than for the wars in Iraq and Peru. The ratio between the two populations' strengths (N_A and N_B) tends to control the slope itself, with greater strength differences resulting in steeper slopes. This suggests that there might be a greater difference between the strengths of the army and insurgency in Colombia than in Iraq or Afghanistan. The total insurgent strength N_A controls the large x roll-off in Fig. 2. Afghanistan and Peru deviate substantially from power laws for large x, which our model interprets as relatively small insurgency strength. Colombia and Iraq hardly deviate from power laws for large x, implying greater insurgency strength.

Because Fig. 3 features data aggregated over size, we replace the detailed grouping dynamics (that is, mechanism (1)) by a time-averaged number of groups. Given the resolution of our data and the typical numbers of observed daily attacks, we take one time-step as equivalent to one day. If a group launches an attack during a day with many other attacks, its media coverage will in general be reduced. If, instead, it launches an attack on a quiet day, its media coverage will increase[20]. Each group receives daily some common but limited information (for example, public radio or newspaper announcements about previous attacks, opposition troop movements, a specific religious holiday, even a shift in weather patterns). The actual content is unimportant provided it becomes the primary input for the group's decision-making process. (See ref. 26 for a full description of an equivalent financial-market version.) Although the groups are heterogeneous in terms of their strategies, they tend to converge towards similar responses when fed the same information[26], thereby generating distributions (black curves) that are almost identical to those observed (green circles). Our model (see Supplementary Information and ref. 26) includes a confidence threshold that must be surpassed before any decision can be made, allowing us to interpret the increase in non-randomness over time for Iraq and Colombia as a decrease in this confidence threshold; that is, the insurgent groups in both wars have become less cautious over time about whether to launch attacks. Reference 30 presents independent empirical evidence that groups of humans do indeed use such generic decision-based mechanisms.

To our knowledge, our model provides the first unified explanation of high-frequency, intra-conflict data across human insurgencies. Other explanations of human insurgency are possible, though any competing theory would also need to replicate the results of Figs 1–3. Our model's specific mechanisms challenge traditional ideas of insurgency based on rigid hierarchies and networks, whereas its striking similarity to multi-agent financial market models[24–26] hints at a possible link between collective human dynamics in violent and non-violent settings[1–19].

METHODS SUMMARY

For the event size distribution (Figs 1 and 2), we use Clauset *et al.*'s methodology[9,10] to estimate power-law exponents, and test power-law and log-normal hypotheses, with the time-aggregated time series of events. This methodology[9,10] is a widely accepted, published state-of-the-art statistical procedure for analysing power-law-like distributions. For the event timing distribution (Fig. 3), we divide the time series for the number of events per day into epochs. These epochs are chosen such that there is no significant trend in the moving-average within each epoch. The precise specification of each epoch's time-window does not affect our main findings. We then generate 10,000 random wars by shuffling the date of the events within each section, averaging across the shuffles. Our model (Fig. 4) replicates the empirical size and timing patterns of Figs 1–3. Full details are given in Methods and Supplementary Information.

Received 5 July; accepted 29 October 2009.

1. Gabaix, X., Parameswaran, G., Plerou, V. & Stanley, H. E. A theory of power law distributions in financial market fluctuations. *Nature* **423**, 267–270 (2003).
2. Lux, T. & Marchesi, M. Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature* **397**, 498–500 (1999).
3. Dussutour, A., Fourcassie, V., Helbing, D. & Deneubourg, J. L. Optimal traffic organization in ants under crowded conditions. *Nature* **428**, 70–73 (2004).
4. Mantegna, R. N. & Stanley, H. E. Scaling behaviour in the dynamics of an economic index. *Nature* **376**, 46–49 (1995).
5. Barabasi, A. L. The origin of bursts and heavy tails in human dynamics. *Nature* **435**, 207–211 (2005).
6. Richardson, L. F. Variation of the frequency of fatal quarrels with magnitude. *J. Am. Stat. Assoc.* **43**, 523–546 (1948).
7. Cederman, L.-E. Modeling the size of wars: from billiard balls to sandpiles. *Am. Polit. Sci. Rev.* **97**, 135–150 (2003).
8. Lim, M., Metzler, R. & Bar-Yam, Y. Global pattern formation and ethnic/cultural violence. *Science* **317**, 1540–1544 (2007).
9. Clauset, A., Shalizi, C. R. & Newman, M. E. J. Power-law distributions in empirical data. *SIAM Rev.* **51**, 661–703 (2009).
10. Clauset, A., Young, M. & Gleditsch, K. S. On the frequency of severe terrorist events. *J. Conflict Resolut.* **51**, 58–87 (2007).
11. Galam, S. & Mauger, A. On reducing terrorism power: a hint from physics. *Physica A* **323**, 695–704 (2003).
12. Johnson, N. *et al.* Universal patterns underlying ongoing wars and terrorism. Preprint at ⟨http://arxiv.org/abs/physics/0605035⟩ (2006).
13. Johnson, D. D. P. & Madin, J. S. in *Natural Security: A Darwinian Approach to a Dangerous World* (eds Sagarin, R. & Taylor, T.) Ch. 11 159–185 (Univ. California Press, 2009).
14. Flack, J. Security in an uncertain world. *Nature* **453**, 451–452 (2008).
15. Adams, E. S. & Mesterton-Gibbons, M. Lanchester's attrition models and fights among social animals. *Behav. Ecol.* **14**, 719–723 (2003).
16. Wrangham, R. W. in *Emerging Synthesis in Science* (ed. Pines, D.) 123–132 (Santa Fe Institute, 1985).
17. Epstein, J. M. *Nonlinear Dynamics, Mathematical Biology, and Social Science* (Addison-Wesley, 1997).
18. Robb, J. *Brave New War: The Next Stage of Terrorism and the End of Globalization* (Wiley, 2007).
19. Hammes, T. *The Sling and the Stone: On War in the 21st Century* (Zenith Press, 2004).
20. Packer, G. Knowing the enemy – can social scientists redefine the war on terror? *New Yorker* 18 December (2006).
21. Ruszczycki, B., Burnett, B., Zhao, Z. & Johnson, N. F. Relating the microscopic rules in coalescence-fragmentation models to the macroscopic cluster-size distribution. *Eur. Phys. J. B* doi:10.1140/epjb/e2009-00354-5 (in the press).
22. May, R. M. & McLean, A. R. *Theoretical Ecology Principles and Applications* 3rd edn (Oxford Univ. Press, 2007).
23. Couzin, I. D., Krause, J., Franks, N. R. & Levin, S. A. Effective leadership and decision making in animal groups on the move. *Nature* **433**, 513–516 (2005).
24. Eguiluz, V. M. & Zimmermann, M. G. Transmission of information and herd behaviour: an application to financial markets. *Phys. Rev. Lett.* **85**, 5659–5662 (2000).
25. D'Hulst, R. & Rodgers, G. J. Exact solution of a model for crowding and information transmission in financial markets. *Int. J. Theor. Appl. Finance* **3**, 609–616 (2000).
26. Johnson, N. F., Jefferies, P. & Hui, P. M. *Financial Market Complexity* (Oxford Univ. Press, 2003).
27. Spirling, A. The next big thing: scale invariance in political science. ⟨http://www.people.fas.harvard.edu/~spirling/documents/powerlawSend.pdf⟩.
28. Palla, G., Barabási, A. L. & Vicsek, T. Quantifying social group evolution. *Nature* **446**, 664–667 (2007).
29. Hicks, M. H.-R. *et al.* The weapons that kill civilians — deaths of children and noncombatants in Iraq, 2003–2008. *N. Engl. J. Med.* **360**, 1585–1588 (2009).
30. Wang, W., Chen, Y. & Huang, J. Heterogeneous preferences, decision-making capacity, and phase transitions in a complex adaptive system. *Proc. Natl Acad. Sci. USA* **106**, 8423–8428 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to J. Restrepo for his involvement in the data collection, and R. Zarama for discussions.

Author Contributions J.C.B., S.G., M.S. and N.F.J. worked on the data and data analysis. S.G., A.R.D. and N.F.J. worked on the model development. All authors participated in the writing and associated discussions, giving detailed feedback in all areas of the project.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to N.F.J. (njohnson@physics.miami.edu).