On the Frequency of Severe Terrorist Events

Aaron Clauset
Santa Fe Institute, New Mexico
and Department of Computer Science
University of New Mexico, Albuquerque

Maxwell Young
Department of Computer Science
University of New Mexico, Albuquerque

Kristian Skrede Gleditsch
Department of Government
University of Essex, United Kingdom
and Centre for the Study of Civil War, Oslo, Norway

In the spirit of Lewis Richardson's original study of the statistics of deadly conflicts, we study the frequency and severity of terrorist attacks worldwide since 1968. We show that these events are uniformly characterized by the phenomenon of "scale invariance," that is, the frequency scales as an inverse power of the severity, $P(x) \propto x^{-\alpha}$. We find that this property is a robust feature of terrorism, persisting when we control for economic development of the target country, the type of weapon used, and even for short time scales. Further, we show that the center of the distribution oscillates slightly with a period of roughly $\tau \approx 13$ years, that there exist significant temporal correlations in the frequency of severe events, and that current models of event incidence cannot account for these variations or the scale invariance property of global terrorism. Finally, we describe a simple toy model for the generation of these statistics and briefly discuss its implications.

Keywords: *terrorism; frequency-severity statistics; scale invariance; Richardson's law*

Richardson (1948) first introduced the concept of "scale invariance"—that is, a power-law scaling between dependent and independent variables—to the study of conflict by examining the frequency of large and small conflicts as a function of their severity. His work demonstrated that for both wars and small-scale homicides, the frequency of an event scales as an inverse power of the event's severity (in this case, the number of casualties). Richardson, and subsequent researchers such as Cederman (2003), have found that the frequency of wars of a size x scales as $P(x) \propto x^{-\alpha}$, where $\alpha \approx 2$ and is called the scaling exponent. Recently, similar power-law statistics have been found to characterize a wide variety of natural phenomena, including disasters such as earthquakes, floods, and

forest fires (Bak and Tang 1989; Malamud, Morein, and Turcotte 1998; Newman 2005), social behavior or organization such as the distribution of city sizes, the number of citations for scientific articles, the number of participants in strikes, and the frequency of words in language (Zipf 1949; Simon 1955; Newman 2005; Biggs 2005), among others. As a reflection of their apparent ubiquity, but somewhat pejoratively, it has even been said that such power-law statistics seem ''more normal than normal'' (Li et al. 2005).

In this article, we extend Richardson's program of study to the most topical kind of conflict—terrorism. Specifically, we empirically study the distributional nature of the frequency and severity of terrorist events worldwide since 1968. Although terrorism as a political tool has a long history (Congleton 2002; Enders and Sandler 2006), it is only in the modern era that small groups of so-motivated individuals have had access to extremely destructive weapons (Shubik 1997; Federal Bureau of Investigation 1999). Access to such weapons has resulted in severe terrorist events such as the August 7, 1998, car bombing in Nairobi, Kenya, which injured or killed over 5,200, and the more well-known attack on September 11, 2001, in New York City, which killed 2,749. Conventional wisdom holds that these rare but severe events are ''outliers''; that is, they are qualitatively different from the more common terrorist attacks that kill or injure only a few people. Although that impression may be true from an operational standpoint, it is false from a statistical standpoint. The frequency-severity statistics of terrorist events are scale invariant, and consequently, there is no fundamental difference between small and large events; both are consistent with a single underlying distribution. This fact indicates that there is no reason to expect that major or more severe terrorist attacks should require qualitatively different explanations than less salient forms of terrorism.

The results of our study are significant for several reasons. First, severe events have a well-documented disproportional effect on the targeted society. Terrorists typically seek publicity, and the media tend to devote significantly more attention to dramatic events that cause a large number of casualties and directly affect the target audience (Wilkinson 1997; Gartner 2004). When governments are uncertain about the strength of their opponents, more severe terrorist attacks can help terrorist groups signal greater resources and resolve and thereby influence a government's response to their actions (Overgaard 1994). Research on the consequences of terrorism, such as its economic impact, likewise tends to find that more severe events exert a much greater impact than less severe incidents (Enders and Sandler 2006, chap. 9). For instance, Navarro and Spencer (2001) report dramatic declines in

share prices on the New York Stock Exchange, NASDAQ, and Amex after the devastating September 11 attacks in the United States. In contrast, although financial markets fell immediately following the July 7, 2005, bombings in London, share prices quickly recovered the next day as it became clear that the bombings had not been as severe as many initially had feared.[1] Recent examples of this non-linear relationship abound, although the tremendous reorganization of the national security apparatus in the United States following the September 11, 2001, attacks is perhaps the most notable in Western society. Second, although researchers have made efforts to develop models that predict the incidence of terrorist attacks, without also predicting the severity, these predictions provide an insufficient guide for policy, risk analysis, and recovery management. In the absence of an accurate understanding of the severity statistics of terrorism, a short-sighted but rational policy would be to assume that every attack will be severe. Later, we will show that when we adapt current models of terrorism to predict event severity, they misleadingly predict a thin-tailed distribution, which would cause us to dramatically underestimate the future casualties and consequences of terrorist attacks. Clearly, we need to better understand how our models can be adapted to more accurately produce the observed patterns in the frequency-severity statistics. That is, an adequate model of terrorism should not only give us indications of where or when events are likely to occur, but also tell us how severe they are likely to be. Toward this end, we describe a toy model that can at least produce the correct severity distribution.

Past research on conflict has tended to focus on large-scale events like wars and to characterize them dichotomously according to their incidence or absence rather than according to their scale or severity. This tendency was recently highlighted by Cederman (2003) for modeling wars and state formation and by Lacina (2006) for civil wars. Additionally, accounting for an event's severity can provide significantly greater guidance to policy makers; for instance, Cioffi-Revilla (1991) accurately predicted the magnitude (the base ten logarithm of total combatant fatalities) of the Persian Gulf War in 1991, which could have helped in estimating the political consequences of the war.

As mentioned above, research on terrorism has also tended to focus on incidence, rather than severity. Recently, however, two of the authors of this study demonstrated for the first time that the relationship between the frequency and severity of terrorist events exhibits the surprising and robust feature of scale invariance (Clauset and Young 2005), just as Richardson (1948) showed for wars. In a subsequent study, Johnson et al. (2005) considered data for fatal attacks or clashes in the guerilla conflicts of Colombia and Iraq, suggesting that these too exhibit scale invariance. Additionally, they claim that the time-varying behavior of these two distributions are trending toward a common power law with parameter $\alpha = 2.5$—a value they note as being similar to the one reported by Clauset and Young (2005) for terrorist events in economically underdeveloped nations. Johnson et al. (2005) then adapted a dynamic equilibrium model of herding behavior on the stock

market to explain the patterns they observed for these guerilla conflicts. From this model, they conjecture that the conflicts of Iraq, Colombia, Afghanistan, Casamance (Senegal), Indonesia, Israel, Northern Ireland, and global terrorism are all converging to a universal distribution with exactly this value of α. We will briefly revisit this idea in a later section. Finally, the recent work of Bogen and Jones (2006) also considers the severity of terrorist attacks primarily via aggregate figures to assess whether there has been an increase in the severity of terrorism over time and to forecast mortality because of terrorism.

This article makes three main contributions. First, we make explicit the utility of using a power-law model of the severity statistics of terrorist attacks and demonstrate the robust empirical fact that these frequency-severity statistics are scale invariant. Second, we demonstrate that distributional analyses of terrorism data can shed considerable light on the subject by revealing new relationships and patterns. And third, we show that, when adapted to predict event severity, existing models of terrorism incidence fail to produce the observed heavy-tail in the severity statistics of terrorism and that new models are needed to connect our existing knowledge about what factors promote or discourage terrorism with our new results on the severity statistics.

Power Laws: A Brief Primer

Before plunging into our analysis, and for the benefit of readers who may be unfamiliar with the topic, we will briefly consider the topics of heavy-tailed statistics and power-law distributions. What distinguishes a power-law distribution from the more familiar normal distribution is its ''heavy tail''; that is, in a power law, there is a nontrivial amount of weight far from the distribution's center. This feature, in turn, implies that events orders of magnitude larger (or smaller) than the mean are relatively common. The latter point is particularly true when compared to a normal distribution, where essentially no weight is far from the mean. Although there are many distributions that exhibit heavy tails, the power law is a particularly special case, being identifiable by a straight line with slope α on doubly logarithmic axes[2], and which appears widely in physics. The power law has the particular form in which multiplication of the argument, for example, by a factor of 2, results in a proportional division of the frequency, for example, by a factor of 4, and the ratio of these values is given by the ''scaling parameter'' α. Because this relationship holds for all values of the power law, the distribution is said to ''scale,'' which implies that there is no qualitative difference between large and small events.

Power-law distributed quantities are actually quite common, although we often do not think of them as being that way. Consider, for instance, the populations of the six hundred largest cities in the United States (from Census 2000). With the average population being only $\langle x \rangle = 165,719$, metropolises like New York City

Figure 1
Frequency-Size Distribution of the 600 Largest U.S. Cities
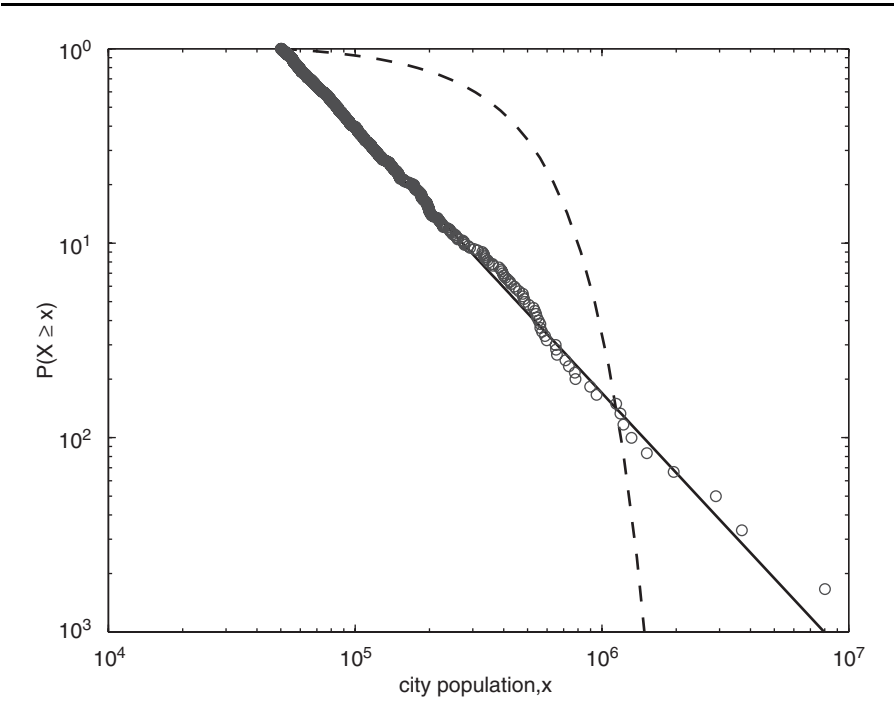
Note: The complementary cumulative distribution function (ccdf) $P(X \geq x)$ of the population x of the six hundred largest cities in the United States, that is, those with $x \geq$ 50,000, based on data from Census 2000. The solid black line shows the power-law behavior that the distribution closely follows, with scaling exponent $\alpha = 2.36(6)$, while the dashed black line shows a truncated normal distribution with the same sample mean. The value in parentheses indicates the uncertainty in our estimate of α.

and Los Angeles would seem to be clear outliers relative to this value. The first clue that this distribution is poorly explained by a truncated normal distribution is that the sample standard deviation $\sigma = 410,730$ is significantly larger than the sample mean. Indeed, if we model the data in this way, we would expect to see 1.8 times fewer cities at least as large as Albuquerque, at 448,607, than we actually do. Furthermore, because it is more than a dozen standard deviations from the mean, we would never expect to see a city as large as New York City, with a population of 8,008,278; for a sample this size, the largest city we would expect to see is Indianapolis at 781,870. Figure 1 shows the actual distribution, plotted on doubly-logarithmic axes, as its complementary cumulative distribution function (ccdf) $P(X \geq x)$, which is the standard way of visualizing this kind of data.[3] The scaling

behavior of this distribution is quite clear, and a power-law model (black line) of its shape is in strong agreement with the data. In contrast, the truncated normal model is a terrible fit.

As a more whimsical second example, consider a world where the heights of Americans were distributed as a power law, with approximately the same average as the true distribution (which is convincingly normal when certain exogenous factors are controlled). In this case, we would expect nearly 60,000 individuals to be as tall as the tallest adult male on record, at 2.72 m. Furthermore, we would expect ridiculous facts such as 10,000 individuals being as tall as an adult male giraffe, one individual as tall as the Empire State Building (381 m), and 180 million diminutive individuals standing a mere 17 cm tall. In fact, this same analogy was recently used to describe the counter-intuitive nature of the extreme inequality in the wealth distribution in the United States (Crook 2006), whose upper tail is also distributed according to a power law.

Although much more could be said about power laws, we hope that the curious reader takes away a few basic facts from this diversion. First, heavy-tailed distributions do not conform to our expectations of a linear or normally distributed world. As such, the average value of a power law is not representative of the entire distribution, and events orders of magnitude larger than the mean are, in fact, relatively common. Second, the scaling property of power laws implies that, at least statistically, there is no qualitative difference between small, medium, and extremely large events, as they are all succinctly described by a very simple statistical relationship. Readers who would like more information about power laws should refer to the extensive review by Newman (2005). With these ideas in hand, we can begin our analysis of the severity statistics of terrorism.

Data Sources for Terrorist Events

Many organizations track terrorist events worldwide, but few provide their data in a form amenable to scientific analysis. The most popular source of information on terrorist events in the political science literature is the ITERATE data set (Mickolus et al. 2004), which focuses exclusively on transnational terrorist events involving actors from at least two countries. In principle, however, and from the standpoint of frequency and severity statistics, we see no reason to restrict our analysis to transnational events. Instead, we use the data contained in the National Memorial Institute for the Prevention of Terrorism (MIPT; 2006) database, which largely overlaps with the ITERATE data but also includes fully domestic terrorist events since at least 1998. We note, however, that our analyses can easily be applied to the portion of the ITERATE data that reports event severity, and indeed, doing so yields evidence similar to that which we present here. Thus, without loss of generality and except where noted, we will focus exclusively on the MIPT data

for the remainder of this article. The MIPT database is itself the compilation of the RAND Terrorism Chronology 1968-1997, the RAND-MIPT Terrorism Incident database (1998-present), the Terrorism Indictment database (University of Arkansas and University of Oklahoma), and DFI International's research on terrorist organizations.

By June 18, 2006, the MIPT database contained records for over 28,445 terrorist events in more than 5,000 cities across 187 countries worldwide since 1968. Although alternative definitions for terrorism exist, the MIPT database uses a relatively standard one that may be summarized as any violent act by a nonstate actor intended to create fear for political purposes. Each entry in the database is quite narrow: it is an attack on a single target in a single location (city) on a single day. For example, the al-Qaeda attacks in the United States on September 11, 2001, appear as three events in the database, one for each of the following locations: New York City, Washington D.C., and Shanksville, Pennsylvania. Each record includes the date, target, city (if applicable), country, type of weapon used, terrorist group responsible (if known), number of deaths (if known), number of injuries (if known), a brief description of the attack, and the source of the information.

Of the nearly 30,000 recorded events, 10,878 of them resulted in at least one person being injured or killed, and we restrict our analyses to these events, since they appear to be the least susceptible to any reporting bias. Furthermore, it is a reasonable assumption that the largest events, because of their severity both in terms of casualties and political repercussions, will have the most accurate casualty estimates. Finally, if there is a systemic bias in the form of a proportional over- or underestimate in the event's severity, it will have only a small effect on the results of our statistical analysis and will not change the core result of scale invariance—as with Richardson's (1948) study of the severity of wars, simply obtaining the correct order of magnitude of an event reveals much of the basic scaling behavior. Throughout the remainder of the article, we take the severity of an event to be either the number of injuries, the number of deaths, or their sum (total casualties), where the severity is always at least one. Unless otherwise noted, we focus exclusively on the statistics of these values.

Frequency-Severity Distributions for Attacks since 1968

Collecting all events since 1968 as a histogram of severities, we show their complementary cumulative distribution functions (ccdf's) $P(X \geq x)$ in Figure 2. The regular scaling in the upper tails of these distributions immediately demonstrates that events orders of magnitude larger than the average event size are not outliers but are instead in concordance with a global pattern in the frequency statistics of terrorist attacks. Significantly, the scaling exists despite large structural and political changes in the international system such as the fall of communism, variations

Figure 2
Frequency-Severity Distributions for Terrorist Attacks Worldwide since 1968
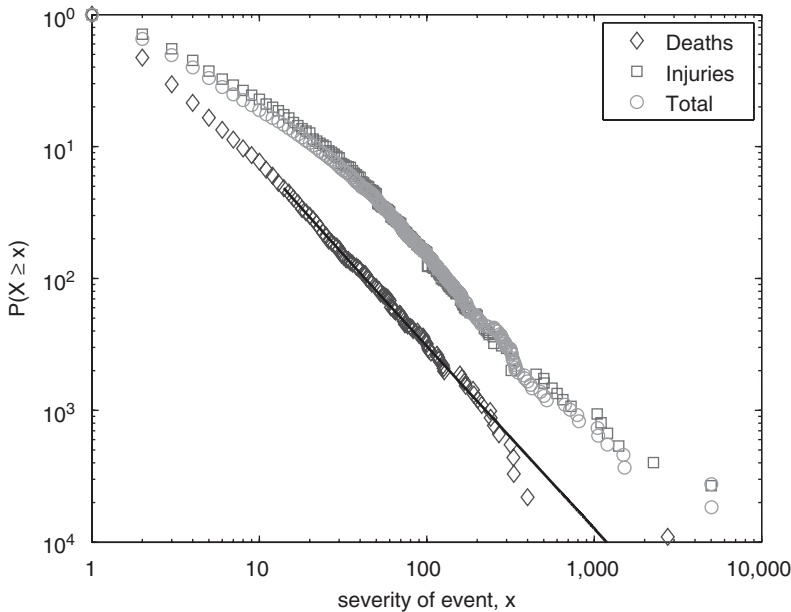
Note: The frequency-severity distributions $P(X \geq x)$ of attacks worldwide since 1968 by injuries, deaths, and their sum. The solid line indicates the power-law scaling found by the maximum likelihood method. Details of fits for these distributions are given in Table 1.

in the type of weapon used, recent developments in technology, the demise of individual terrorist organizations, and the geographic distribution of events themselves. In subsequent sections, we will examine the robustness of the scale invariance property to both categorical and temporal analysis.

If we make the idealization that events are independent and identically distributed (iid), we may model the distribution as a power law with some exponent α, where the scaling behavior holds only for values at least as large as some lower-bound x_{min}. Obviously, significant correlations exist between many terrorist events, and such an idealization is made only for the purpose of doing a distributional analysis. Using the method of maximum likelihood, we estimate two parameters of the power-law model from the data (details of our statistical methodology are discussed in the appendix). Models found in this way for the full distributions described above are summarized in Table 1. Using the Kolmogorov-Smirnov goodness-of-fit test, we find that these simple iid models are a surprisingly good representation of

Table 1
Summary of Distributions and Models Shown in Figure 2

Distribution	N	$\langle X \rangle$	σ_{std}	X_{max}	N_{tail}	α	X_{min}	$p_{KS} \geq$
Injuries	7,456	12.77	94.45	5,000	259	2.46(9)	55	0.41
Deaths	9,101	4.35	31.58	2,749	547	2.38(6)	12	0.94
Total	10,878	11.80	93.46	5,213	478	2.48(7)	47	0.99

Note: A summary of the distributions shown in Figure 2, with power-law fits from the maximum likelihood method. N (N_{tail}) depicts the number of events in the full (tail) distribution. The parenthetical value depicts the standard error of the last digit of the estimated scaling exponent.

the death and total severity distributions (both $p_{KS} > 0.9$), although a more marginal representation of the injuries distribution ($p_{KS} > 0.4$).[4]

In a later section we will see that we can further decompose these distributions into their components, each of which is strongly scale invariant but with different scaling and limit parameters. As mentioned earlier, the power law is not the only distribution with a heavy tail, and although testing all such alternatives is beyond the scope of this article, we considered another common distribution, the log-normal (see, e.g., Serfling 2002), and found in all cases that we may convincingly reject this model ($p_{KS} < 0.05$).

Evolution of Terrorism over Time

Because events in the database are annotated with their incidence date, we may write them down as a time series and investigate the severity distribution's behavior as a function of time.[5] Although we are ultimately interested in the property of scale invariance over time, we first consider a simple, model-agnostic measure of the distribution's shape: the average log-severity. Sliding a window of twenty-four months over the 38.5 years of event data, we compute the average log-severity (deaths) of events within each window.

For highly skewed distributions, such as those we show in Figure 2, the average log-severity measures the position on the independent axis of the distribution's center. The average log-severity is significantly less sensitive to variations in the length of the upper tail, which may arise from the occasional presence of rare but severe events, than is the average severity. The resulting time series of this measure is shown in the upper pane of Figure 3, along with one standard deviation. Notably, this function is largely stable over the nearly forty years of data in the MIPT database, illustrating that the center of the distribution has not varied substantially over that time.

A closer examination of the fluctuations, however, suggests the presence of potential periodic variation. We investigate this possibility by taking the

Figure 3
Average Log-Severity of Events over Time, and Its Autocorrelation Function
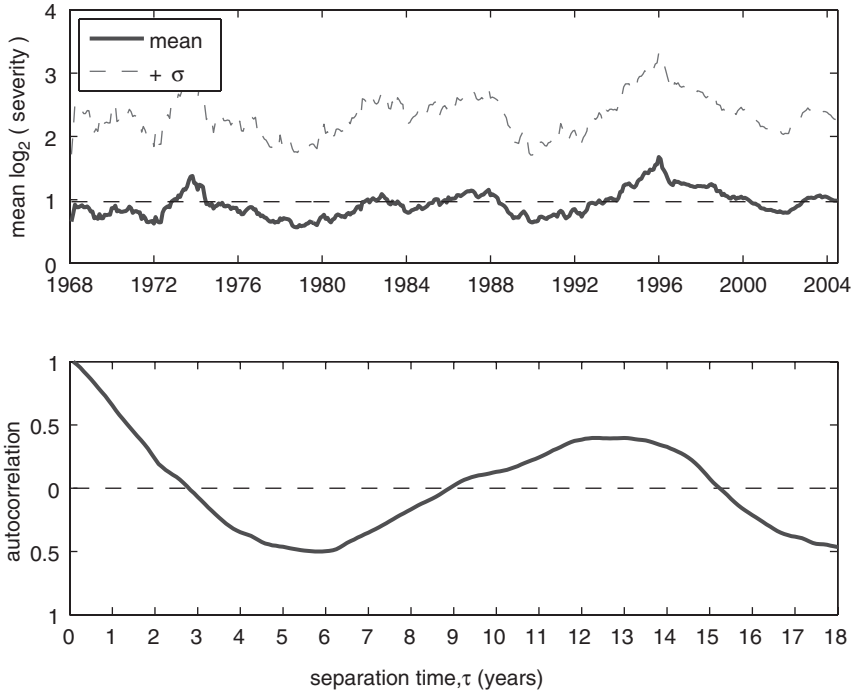
Note: (upper) The average log-severity (deaths) of events within a sliding window of twenty-four months, for the entire 38.5 years of data. The upper dashed line indicates one standard deviation, while the other shows the average log-severity for the entire span of time. (lower) The autocorrelation function of the average log-severity, illustrating a strong periodicity in the breadth of the distribution at roughly $\tau \approx 13$ years. Similar results apply when we analyze total or injury severity but with slight changes to the magnitude or location of the anomalous peak in the autocorrelation function.

autocorrelation function (ACF) of the time series, which we show in the lower-pane. The noticeable sinusoidal shape in the ACF shows that the fluctuations do exhibit a strong degree of periodicity on the order of $\tau \approx 13$ years. If we vary the size of the window, for example, windows between twelve and sixty or more months (data not shown), the location and magnitude of the peak are, in fact, quite stable. But these features do vary slightly if we instead either examine the total or injury distributions or truncate the time series. As such, we conjecture that some periodicity is a natural feature of global terrorism, although we have no explanation for its origin. It has been suggested that the $\tau \approx 13$ value may be related to the

modal life expectancy of the average terrorist group. However, we caution against such conclusions for now, since these aforementioned variations on our analysis can shift the peak by several years.

## Scale Invariance over Time

Turning now to the question of scale invariance over time, we again use a sliding window of two years but now shifted forward by one year at a time. To remain parsimonious, we make the idealization that events within each window were drawn iid from a two-parameter power-law model. After fitting such a model to each window's frequency-severity distribution, we calculate its statistical significance as a way to check the model's plausibility for that time period. Obviously, this assumption of no temporal correlations is quite strong, and, where appropriate, we discuss what light our analysis sheds on its accuracy. Johnson et al. (2005) used a similar approach to study the time-varying distributions for the conflicts in Colombia and Iraq but did not consider the accuracy of their models' fit or give any measure of their statistical significance.

In Figure 4a, we show the estimated scaling parameters α for each time period. For the first thirty years of data, the scaling parameter appears to fluctuate around $\alpha \approx 2$, which suggests that the scaling behavior was relatively stable over this period. Subsequent to 1998, when a larger number of domestic events were incorporated into the database, the scaling parameter shifts upward to $\alpha \approx 2.5$ but again shows no consistent trend in any direction. This shift, taken with the apparent stability of the scaling behavior over time, suggests that the absence of domestic events before 1998 may have biased those distributions toward more shallow scaling; that is, before 1998 larger events appear to be more common.

Although many (41 percent) of these iid power-law models appear to match the distribution of severities quite well ($p_{KS} > 0.9$), nearly half (49 percent) achieve only a middling level of statistical significance ($0.5 < p_{KS} \leq 0.9$; Figure 4b, upper). That is, there are significant temporal correlations within the time series, or perhaps there are strong but temporally localized deviations from the long-term structure of the power-law distribution, which cause our simple model to yield a poor fit at these times. Either case is unsurprising for this kind of real-world data. An interesting line of future inquiry would be a close study of the tail events' political context, which may reveal the origin of their correlations and explain when temporally local deviations from the long-term behavior occurred.

Furthermore, we observe that the frequency of the most severe events, that is, events in the upper tail of the distribution, has not changed much over the past thirty years. In Figure 4b (lower pane), we plot the reciprocal of those frequencies, the mean interevent intervals, for each two-year period. Notably, from 1977 to 1997, the interevent interval for extreme events averaged 6.9 ± 3.7 days, while

Figure 4

Estimated Parameters of Power-Law Models over Time, 1968 to Present
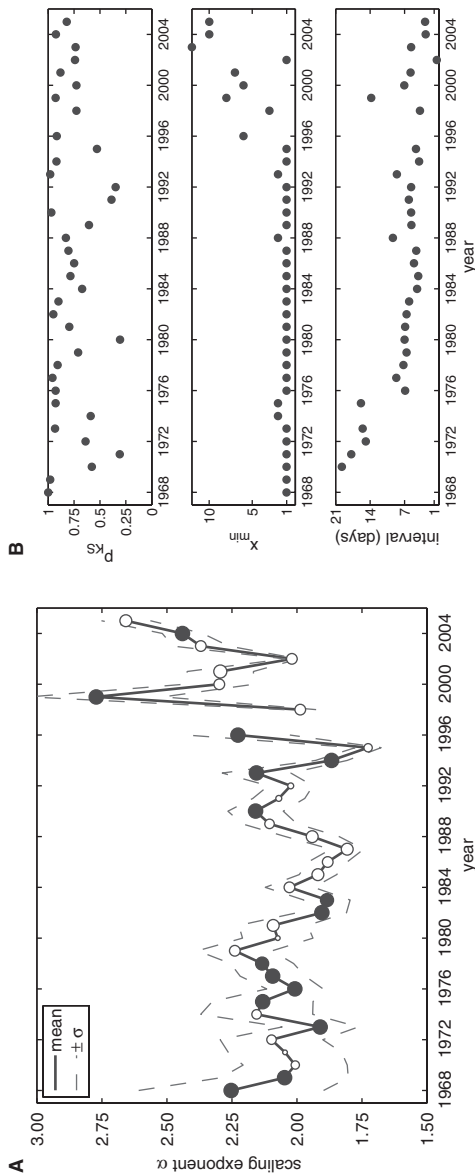
Note: Results of fitting a simple power-law model to the tail of the severity (deaths) distribution of terrorist events within discrete periods of time since 1968. We divide time into two-year periods, sliding the window forward one year at a time; similar results apply for larger windows. (a) The average scaling exponent α for each two-year period, with circle size being proportional to the statistical significance value p_{KS}; solid circles indicate $p > 0.9$. We omit the data point for 1997, since it spans the transition of the database management. (b) Three panes showing aspects of the models; (top) significance values computed from a one-sided Kolmogorov-Smirnov test, showing that most models do not achieve high significance, (middle) the estimated x_{min} values, and (bottom) the average interevent interval for events in the tail, that is, those with severity greater than x_{min}.

from 1998 to 2006, it averaged 5.3 ± 4.0 days. Although this result may appear to contradict recent official reports that the frequency of terrorist attacks worldwide has increased dramatically in the past few decades (United States Department of State 2003) or that the frequency of major events has decreased, it does not. Instead, the situation is slightly more complicated: our analysis suggests that the changes in event frequencies have not been evenly distributed with respect to their severity but rather that less severe attacks are now relatively more frequent, while the frequency of major or tail-events has remained unchanged. This behavior is directly observable as the upward movement of the lower-bound on the scaling region in recent years, precisely when attacks overall are thought to be more frequent (Figure 4b, middle pane).

Taking the above results together with those of the average log-severity time-series (Figure 3) in the previous section, we can reasonably conclude that the dominant features of the frequency-severity statistics of terrorism have not changed substantially over the past 38.5 years. That is, had some fundamental characteristic of terrorism changed in the recent past, as we might imagine given recent political events, the frequency-severity distribution would not display the degree of stability we observe in these statistical experiments.

Variation in Scale Invariance by Target-Country Industrialization

Returning to the full distributions, we now consider the impact of industrialization on the frequency-severity statistics—given that each attack is executed within a specific country, we may ask whether there is a significant difference in the scaling behaviors of events within industrialized and nonindustrialized countries. Toward this end, we divide the events since 1968 into those that occurred within the thirty Organisation for Economic Co-operation and Development (OECD) nations (1,244 events, or 11 percent) and those that occurred throughout the rest of the world (9,634 events, or 89 percent). We plot the corresponding total severity distributions in Figure 5a, and give their summary statistics in Table 2.

Most notably, we find substantial differences in the scaling of the two distributions, where industrialized-nation events scale as $\alpha_{OECD} = 2.02(9)$, while nonindustrialized-nation events scale more steeply, as $\alpha_{non-OECD} = 2.51(7)$. That is, while events have been, to date, less likely to occur within the major industrialized nations, when they do, they tend to be more severe than in nonindustrialized nations. Although this distinction is plausibly the result of technological differences, that is, industrialization itself makes possible more severe events, it may also arise because industrialized nations are targeted by more severe attacks for political reasons. For instance, the OECD events are not uniformly distributed over the thirty OECD nations but are disproportionately located in eight states: Turkey

Table 2
Summary of Distributions and Models Shown in Figure 5

Distribution	N	$\langle X \rangle$	σ_{std}	X_{max}	N_{tail}	α	X_{min}	$p_{KS} \geq$
OECD[a]	1,244	17.65	206.28	5,012	158	2.02(9)	13	0.61
Non-OECD	9,634	11.04	66.09	5,213	438	2.51(7)	47	0.84
Chem/bio	19	274.11	1,147.48	5,012	19	1.5(2)	1	0.89
Explosives	4,869	18.93	90.61	5,213	412	2.52(7)	49	0.60
Fire	133	16.79	107.14	1,200	85	1.9(1)	2	0.99
Firearms	4,603	4.09	24.52	1,058	744	2.37(5)	5	0.92
Knives	254	2.43	7.01	107	52	2.6(2)	3	0.99
Other	1,000	9.30	158.79	5,010	189	2.17(9)	5	0.99

Note: A summary of the distributions shown in Figure 5, with power-law fits from the maximum likelihood method. N (N_{tail}) depicts the number of events in the full (tail) distribution. The parenthetical value depicts the standard error of the last digit of the estimated scaling exponent. As described in the text, the statistical significance of the explosives distribution model increases to $p_{KS} \geq 0.82$ when we control for suicide explosive attacks.
a. OECD = Organisation for Economic Co-operation and Development.

(335 events), France (201), Spain (109), Germany (98), the United States of America (93), Greece (76), Italy (73), and the United Kingdom (62). These eight account for 84.2 percent (1,047) of all such OECD events, and 141 of those are tail events; that is, their total severity is at least $x_{min} = 13$, the estimated lower limit of the tail region. These eight nations account for 89.2 percent of the severe events, suggesting that industrialization alone is a weak explanation of the location of severe attacks and that political factors must be important.

Variation in Scale Invariance by Weapon Type

As our final characterization of the frequency-severity distribution's scale-invariance, we consider the connection between technology, represented by the type of weapon used in an attack, and the severity of the event. Figure 5b shows the total severity distributions for chemical or biological weapons, explosives (including remotely detonated devices), fire, firearms, knives, and a catchall category, other (which also includes unconventional[6] and unknown weapons). We find that these component distributions themselves exhibit scale invariance, each with a unique exponent α and lower limit of the power-law scaling x_{min}. However, for the chemical or biological weapons and the explosives distributions, we must make a few caveats. In the former case, the sparsity of the data reduces the statistical power of our fit, and as discussed by Bogen and Jones (2006), the severity of the largest such event, the 1998 sarin gas attack in Tokyo, is erroneously high. For the latter distribution, another phenomenon must govern the shape of the lower tail, and we

Figure 5
Frequency-Severity Distributions by Industrialization and by Weapon-Type
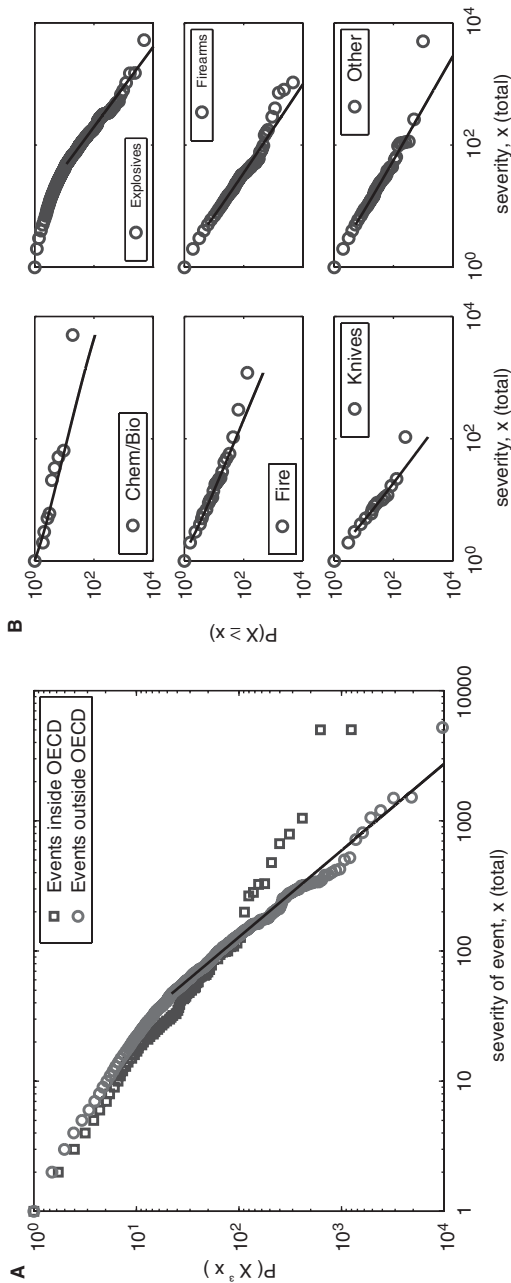
Note: (a) The frequency-severity (total) distributions $P(X \geq x)$ of attacks worldwide between February 1968 and June 2006, divided among nations inside and outside of the OECD. (b) Total-severity distributions for six weapon types: chemical or biological agents (0.2 percent of events), explosives (44.8 percent; includes remotely detonated devices), fire (1.2 percent), firearms (42.3 percent), knives (2.3 percent), and other (9.2 percent; includes unconventional and unknown weapon types). For both figures, solid lines indicate the fits, described in Table 2, found by the maximum likelihood method.

72

investigate its causes below. Table 2 summarizes the distributions and their power-law models.

By partitioning events by weapon type, we now see that the origin of the bending in the lower tail of the injury and total severity distributions (Figure 2a) are primarily because of explosive attacks; that is, there is something about attacks utilizing explosives that makes them significantly more likely to injure a moderate or large number of people than other kinds of weapons. However, this property fails for larger events, and the regular scaling resumes in the upper tail. In contrast, we see no such change in the scaling behavior in the lower tail for other weapons—this demonstrates that the property of scale invariance is largely independent of the choice of weapon. Furthermore, by partitioning events according to their weapon type, we retain high estimates of statistical significance ($p_{KS} > .9$).

What property of explosives attacks can explain the large displacement of the upper tail in that distribution? Pape (2003) demonstrated through a careful analysis of all suicide attacks between 1980 and 2001 that suicide attacks cause significantly more deaths than nonsuicide attacks on average, being 13 and 1, respectively. Similarly, for our data set, the average total severity for suicide attacks using explosives is 41.11, while nonsuicide attacks have an average total severity of 14.41. Controlling for these attacks (692 events, or 12.9 percent) does not significantly change the curvature of the lower tail in the explosives distribution. It does, however, improve the statistical significance of our best-fit model to the upper tail ($\alpha = 2.55(9)$, $x_{min} = 47$, $p_{KS} \geq .82$), suggesting that the severity of suicide explosives attacks deviates strongly from the general scaling behavior and furthermore that such attacks are not the source of the lower tail's curvature. Conditioning on additional factors, either singly or jointly, such as the target, tactic, or geographic region, can reduce the curvature in the lower tail to varying degrees but can never eliminate it (results not shown).

By analyzing the sequence of events, however, we find evidence that the curvature is at least partially a temporal phenomenon. When we divide events into the four decades beginning with 1968, 1978, 1988, and 1998, we see that the location of the upper tail x_{min} increases over time, ranging from between 2 and 20 for the first three decades to 49 for the most recent decade. Indeed, because most of the explosives events in the database occurred recently (3,034 nonsuicide events or 72.6 percent), the scaling behavior of this decade dominates the corresponding distribution in Figure 5b. Separating the data by time, however, yields more statistically significant models, with $p_{KS} \geq .8$ for the latter three decades and progressively more curvature in the lower tail over time. Thus, we cannot wholly attribute the curvature to the inclusion of domestic events in more recent years, although certainly it is largest then. Rather, its behavior may be a function of changes in the explosives technology used in terrorist attacks over the past forty years. The validation of this hypothesis, however, is beyond the scope of the current study, and we leave it for future work.

A Regression Model for the Severity of Terrorist Events

There is an extensive literature on what factors promote terrorism and make governments more likely to become targets of terrorism. We refer to Reich (1990), Pape (2003), and Rosendorff and Sandler (2005) for overviews of existing studies of terrorism. Notably, however, existing studies say nothing about the frequency-severity distribution of events, and empirical research on terrorism has tended to focus on predicting attack incidence. In this section, we consider to what extent models proposed to predict the incidence of terrorism data can account for the severity of terrorism and to what extent they can reproduce the observed frequency-severity distribution.

As a recent example of empirical studies on the frequency of terrorist attacks, we use that of Li (2005). Although different studies have suggested different features to predict variation in terrorist incidents, the Li study is both careful and generally representative of the structure of cross-country comparative studies. Li empirically explores the impact of a large number of political and economic factors that have been hypothesized to make transnational terrorist incidents more or less likely, and argues that while some features of democratic institutions, such as greater executive constraints, tend to make terrorist incidents more likely, other features, such as democratic participation, are associated with fewer incidents. Model (1) in Table 3 displays the coefficient estimates for Li's original results from a negative binomial regression of the number of transnational terrorist events, with each country-year as the unit of observation. We refer to the original Li (2005) article for all details on variable construction, and so forth.

Since our data are based on terrorist incidents that are not limited to transnational events, we first replicate the Li model for incidents in the MIPT (2006) data to ensure that our results are not an artifact of systematic differences between transnational-only and transnational-plus-domestic terrorist events. The coefficient estimates for the Li model applied to the number of incidents in the MIPT data shown as model (2) in Table 3 are, in general, reasonably similar to the results for the original model (1), suggesting that the model behaves similarly when applied to the two sources of data on terrorism.

Next, we examine to what extent the right-hand-side covariates in the Li model allow us to predict differences in the severity of terrorism. Model (3) in Table 3 displays the results for a negative binomial regression of the number of deaths among the lethal events in the MIPT data. Comparing the size of the coefficient estimates to their standard errors suggests that none of these coefficients are distinguishable from 0, with the possible exception of the estimate for Europe and the post–cold war period. In other words, none of the factors proposed by Li seem to be good predictors of the severity of terrorist events. Moreover, the proposed Li model fails to generate predictions that in any way resemble the observed variation in the number of deaths: the largest predicted number of deaths for any observation

Table 3
Regression Models for Event Incidence and Severity

Variable	(1) No. Attacks (ITERATE)	(2) No. Attacks (MIPT)	(3) Deaths by Event	(4) Deaths by Country-Year
Government constraint	0.061	0.102	−0.013	0.046
	(0.023)	(0.030)	(0.013)	(0.038)
Democratic participation	−0.009	−0.007	−0.001	−0.011
	(0.004)	(0.006)	(0.003)	(0.007)
Income inequality	0.001	−0.001	0.003	−0.002
	(0.014)	(0.016)	(0.007)	(0.021)
Per capita income	−0.177	−0.161	0.008	−0.222
	(0.11)	(0.14)	(0.047)	(0.15)
Regime durability	−0.076	−0.109	0.039	0.010
	(0.047)	(0.060)	(0.024)	(0.067)
Size	0.118	0.0494	−0.014	−0.001
	(0.044)	(0.054)	(0.015)	(0.079)
Government capability	0.275	0.189	−0.018	0.072
	(0.14)	(0.18)	(0.061)	(0.21)
Past incident	0.547	0.717	−0.009	0.789
	(0.045)	(0.052)	(0.024)	(0.081)
Post −cold war	−0.578	−0.253	0.104	−0.036
	(0.097)	(0.11)	(0.061)	(0.16)
Conflict	−0.170	−0.046	0.294	0.072
	(0.11)	(0.13)	(0.13)	(0.39)
Europe	0.221	−0.263	−0.133	−0.589
	(0.20)	(0.34)	(0.075)	(0.49)
Asia	−0.494	−0.684	0.239	−0.542
	(0.25)	(0.28)	(0.13)	(0.36)
America	−0.349	−0.681	−0.098	−1.125
	(0.15)	(0.23)	(0.073)	(0.30)
Africa	−0.423	−0.462	0.022	−0.538
	(0.18)	(0.21)	(0.12)	(0.31)
Constant	−0.443	0.805	1.591	2.548
	(1.54)	(1.89)	(0.65)	(2.63)
N	2,232	2,232	1,109	2,232
Log-likelihood	−3,805.791	−3,300.011	−2,897.375	−2,268.129
LR $-\chi^2$	1,151.842	507.427	151.709	373.293

Note: Coefficients for a negative binomial regression model on terrorist event incidence, after Li (2005), and its ability to predict observed severity statistics; parenthetical entries give robust standard errors.

in the observed sample is less than 10, far below the actual observed maximum of 2,749 (i.e., the September 11, 2001, attack on the World Trade Center).

The original Li model examines the number of incidents by country-year, and it may therefore be argued that looking only at events with casualties could

understate the possible success of the model in identifying countries that are unlikely to become targets of terrorist incidents. The results for the Li model applied to the total events for all country-years, model (4) in Table 3, however, do not lend much support to this idea. Very few of the features emphasized by Li have coefficient estimates distinguishable from 0 by conventional significance criteria, and the highest predicted number of deaths for any one country-year in the sample is still less than 16. As such, this model is clearly not able to generate the upper tail of the observed frequency-severity distribution.

A Toy Model for Scale Invariance through Competitive Forces

Having shown that a representative model of terrorism incidence is a poor predictor of event severity, we now consider an alternative mechanism by which we can explain the robust statistical feature of scale invariance. As it turns out, power-law distributions can arise from a wide variety of processes (Kleiber and Kotz 2003; Mitzenmacher 2004; Newman 2005; Farmer and Geanakoplos 2006). In the case of disasters such as earthquakes, floods, forest fires, strikes, and wars, the model of self-organized criticality (Bak, Tang, and Wiesenfeld 1987), a physics model for equilibrium critical phenomena[7] in spatially extended systems, appears to be the most reasonable explanation (Bak and Tang 1989; Malamud, Morein, and Turcotte 1998; Cederman 2003; Biggs 2005), since events themselves are inherently spatial. However, such models seem ill-suited for terrorism, where the severity of an event is not merely a function of the size of the explosion or fire. That is, the number of casualties from a terrorist attack is also a function of the density of people at the time and location of the attack and of the particular application of its destructive power; for example, a small explosion on an airplane can be more deadly than a large explosion on solid ground.[8]

In the context of guerilla conflicts, Johnson et al. (2005; 2006) have adapted a dynamic equilibrium model of herding behavior on the stock market to produce frequency-severity distributions with exponents in the range of 1.5 to 3.5, depending on a parameter that is related to the rates of fragmentation and coalescence of the insurgent groups; they conjecture that the value 2.5, in particular, is universal for all asymmetric conflict, including terrorism. Given the variation in the scaling behaviors that we measure for different aspects of terrorism (Figures 2, 4a, 5a, and 5b), this kind of universalism may be unwarranted. As an alternative explanation to the origin of the scale invariance for terrorism, we propose and analyze a simple, nonspatially extended toy model of a stochastic, competitive process between states and nonstate actors (Clauset and Young 2005). The model itself is a variation of one described by Reed and Hughes (2002) and can produce exponents that vary depending on the choice of model parameters—a feature necessary to explain the

different scaling behaviors for industrialization and weapon types. Central to our model are two idealizations: that the potential severity of an event is a certain function of the amount of planning required to execute it and that the competition between states and nonstate actors is best modeled by a selection mechanism in which the probability that an event is actually executed is inversely related to the amount of planning required to execute it.

Consider a nonstate actor (i.e., a terrorist) who is planning an attack. Although the severity of the event is likely to be roughly determined before planning begins, we make the idealization that the potential severity of the event grows with time, up to some finite limit imposed perhaps by the choice of weapon (as suggested by Figure 5), the choice of target, or the availability of resources. If we further assume that the payoff rate on additional planning is proportional to the amount of time already invested, that is, increasing the severity of a well-planned event is easier than for a more ad hoc event, then the potential severity of the event can be expressed as $P(t) \propto e^{\kappa t}$, where $\kappa > 0$ is a constant.

However, planned events are often prevented, aborted, or executed prematurely possibly as a result of intervention by a state. This process by which some events are carried out while others are not can be modeled as a selection mechanism. Assuming that the probability of a successful execution is exponentially related to the amount of time invested in its planning, perhaps because there is a small chance at each major step of the planning process that the actors will be incarcerated or killed by the state or will abandon their efforts, we can relate the severity of a real event to the planning time of a potential event by $x \propto e^{\lambda t}$, where $\lambda < 0$ is a constant. Thus, to derive the distribution of real event severities, after the selection mechanism has filtered out those events that never become real, we must solve the following identity from probability theory:

$$\int P(x)\,dx = \int P(t)\,dt.$$

Doing so yields $P(x) \propto x^{-\alpha}$, where $\alpha = 1 - \kappa/\lambda$. Again considering the competitive nature of this process, it may be plausible that states and actors will, through interactions much like the coevolution of parasites and hosts, develop roughly equal capabilities on average but perhaps with a slight advantage toward the state by virtue of its longevity relative to terrorist organizations, such that $|\kappa| \gtrsim |\lambda|$. In this case, we have a power law with exponent $\alpha \gtrsim 2$, in approximate agreement with much of our empirical data.

Although our toy model makes several unrealistic idealizations, its foundational assumptions fit well with the modern understanding of terrorism and also with examples of recent attacks and foiled attempts. Whereas the plans for September 11, 2001, attacks in the United States are believed to have been under way since 1996,[9] subsequent attacks and attempts in the United Kingdom carried out by less

organized groups and with less advance planning have failed to create a similar impact. For example, the July 21, 2005, attacks on the London Underground are now believed to have been a direct copycat effort initiated after the prior July 7 bombings. The attack was spectacularly unsuccessful: none of the four bombs' main explosive charges actually detonated, and the only reported casualty at the time was later found to have died from an asthma attack. Even though the suspects initially managed to flee, all were later apprehended.

The competitive relationship of states and nonstate actors has been explored in a variety of other contexts. Hoffman (1999) suggests that the state's counterterrorism efforts serve as a selective measure, capturing or killing those actors who fail to learn from their peers' or predecessors' mistakes, leaving at large the most successful actors to execute future attacks. Overgaard (1994), Sandler and Arce M. (2003), Sandler and Lapan (1988), and Arce M. and Sandler (2005) give a similar view, arguing that the actions of states and actors are highly interdependent—that actors typically make decisions on who, where, when, or what to attack based on a careful assessment of the likelihood and impact of success, with these factors being intimately related to the decisions states make to discourage certain forms of attacks or responses. Governments make a similar calculus, although theirs is primarily reactive rather than proactive (Arce M. and Sandler 2005). Looking forward, a game theoretic approach such as the one used by Sandler and Arce M. (2003) to produce practical counterterrorism policy suggestions will likely be necessary to capture this interdependence, although presumably it will be roughly similar to the selective process we describe above.

Obviously, the practical, geopolitical, and cultural factors relevant to a specific terrorist attack are extremely complex. Although our toy model intentionally omits them, they presumably influence the values assumed by the model parameters and are essential for explaining the variety of scaling exponents we observe in the data, for example, the different scaling exponents for OECD and non-OECD nations and for attacks perpetrated using different weapons. It may be possible to incorporate these factors by using a regression approach to instead estimate the parameter values of our toy model rather than to directly estimate the event severity.

Discussion and Conclusions

Many of the traditional analyses of trends in terrorism are comparative, descriptive, historical, or institutional, and those that are statistical rely on assumptions of normality and thus treat rare but severe events as qualitatively different from less severe but common events (Reich 1990; Federal Bureau of Investigation 1999; United States Department of State 2003; Rosendorff and Sandler 2005). By demonstrating that Richardson's (1948) discovery of scale invariance in the frequency-severity statistics of wars extends to the severity statistics of terrorism, we show

that these assumptions are fundamentally false. Our estimates of the scaling behavior for terrorism, however, differ substantially from that of the severity of wars; in the latter case, the frequency-severity distribution scales quite slowly, with $\alpha_{war} = 1.80(9)$, while the distribution scales much more steeply for terrorism, $\alpha_{deaths} = 2.38(6)$, indicating that severe events are relatively less common in global terrorism than in interstate warfare.

Taking Richardson's program of study on the statistics of deadly human conflicts together with the extensive results we discuss here, our previous, preliminary study of terrorism (Clauset and Young 2005), and the study by Johnson et al. (2005; 2006) of insurgent conflicts, we conjecture first that scale invariance is a generic feature of the severity distribution of all deadly human conflicts, and second that it is the differences in the type of conflict that determine the particular scaling behavior, that is, the values of the scaling exponent α and the lower limit of the scaling $x_{\min}$. Indeed, this variation is precisely what we observe when we control for attributes such as the degree of economic development and the type of weapon used in the attack. In honor of Richardson and his pioneering interest in the statistics of deadly conflict, we call our conjecture Richardson's Law. A significant open question for future work remains to determine how and why the distinguishing attributes of a conflict, such as the degree of asymmetry, the length of the campaign, and the political agenda, and so forth, affect the observed scaling behavior.

With regard to counterterrorism policy, the results we describe here have several important implications. First, the robustness of the scale invariant relationship between the frequency and severity of attacks demonstrates the fact that severe events are not fundamentally different from less severe ones. As such, policies for risk analysis and contingency planning should reflect this empirical fact. Second, although severe events do occur with much greater frequency than we would expect from our traditional thin-tailed models of event severity, their incidence has also been surprisingly stable over the past thirty years (Figure 4b, lower pane). This point suggests that from an operational standpoint and with respect to their frequency and severity, there is nothing fundamentally new about recent terrorist activities worldwide. Third, limiting access to certain kinds of weapons and targets is clearly important, with this being particularly true for those that are inherently more likely to produce a severe event, such as high explosives or targets like airplanes and other mass transit systems. But severe events themselves are not only associated with one or a few weapon types (or targets). Restricting access to some weapons and targets will likely induce the substitution of less easily restricted ones (Enders and Sandler 2006)—a contingency for which we should plan. Fourth, the trend we identify for explosives, that is, that such attacks have produced progressively more casualties over time, is particularly distressing given the sheer number of explosives attacks in the recent past. Both their severity and their popularity

suggest that current international regulation of explosives technology is failing to keep these weapons out of the hands of terrorists and that current diplomacy is failing to keep terrorists from resorting to their use. And finally, although it may be tempting to draw an analogy between terrorism and natural disasters, many of which also follow power-law statistics, we caution against such an interpretation. Rather, a clear understanding of the political and socioeconomic factors that encourage terrorist activities and an appropriate set of policies that directly target these factors may fundamentally change the frequency-severity statistics in the future and break the statistical robustness of the patterns we have observed to date.

In closing, the discovery that the frequency of severe terrorist attacks follows a robust empirical law opens many new questions and points to important gaps in our current understanding of both the causes and consequences of terrorism. Although we have begun to address a few of these, such as showing that the severity of suicide attacks using explosives does not follow the same frequency-severity statistics as other forms of terrorism, many more remain. We hope to see the community of conflict researchers making greater use of these new ideas in future research on terrorism.

Appendix
Statistical Methodology, and the
Use of Power Laws in Empirical Studies

Because the use of power laws and other heavy-tailed distributions in the social sciences is a relatively new phenomenon, the statistical tools and their relevant characteristics may not be familiar to some readers. This appendix thus serves to both explain our statistical methodology and to give the interested reader a brief tutorial on the subject. We hope that this material illuminates a few of the subtleties involved in using power laws in real-world situations. Readers interested in still more information should additionally refer to Newman (2005) and Goldstein, Morris, and Yen (2004).

To begin, we note that there are two distinct kinds of power laws, a real-valued or continuous kind and a discrete kind. Although both forms have many characteristics in common, the numerical methods one employs in empirical studies can be quite different, depending on whether the data are best treated as continuous or discrete. Examples of the former might be voltages on power lines, the intensity of solar flares, or the magnitude of earthquakes. In cases where discrete data takes values that are quite large, they can often be safely treated as if they were continuous variables, such as for the population of U.S. cities, book sales in the United States, or the net worth of Americans. In the social sciences, however, data more frequently assume integer values where the maximum value is only a few orders of

magnitude larger than the minimum; that is, the tail is heavy but rather short. Examples of this kind of data might be the number of connections per person in a social network, casualty statistics for terrorist attacks, and word frequencies in a text. If such data are treated as a continuous variable, estimates of the scaling behavior or other statistical analyses can be significantly biased.

Instead, these heavy but relatively short tails should be modeled explicitly as a discrete power law,

$$P(x) = x^{-\alpha}/\zeta(\alpha)$$

with x assuming only integer values greater than zero, and $\zeta(\alpha)$ being the Riemann zeta function, the normalization constant. In what follows, we will first consider the necessity of generating a random deviate with a power-law distribution and then consider methods for estimating power-law parameters from data itself. Both sections describe the statistical methods employed in this study, and provide a brief comparison with alternative methods.

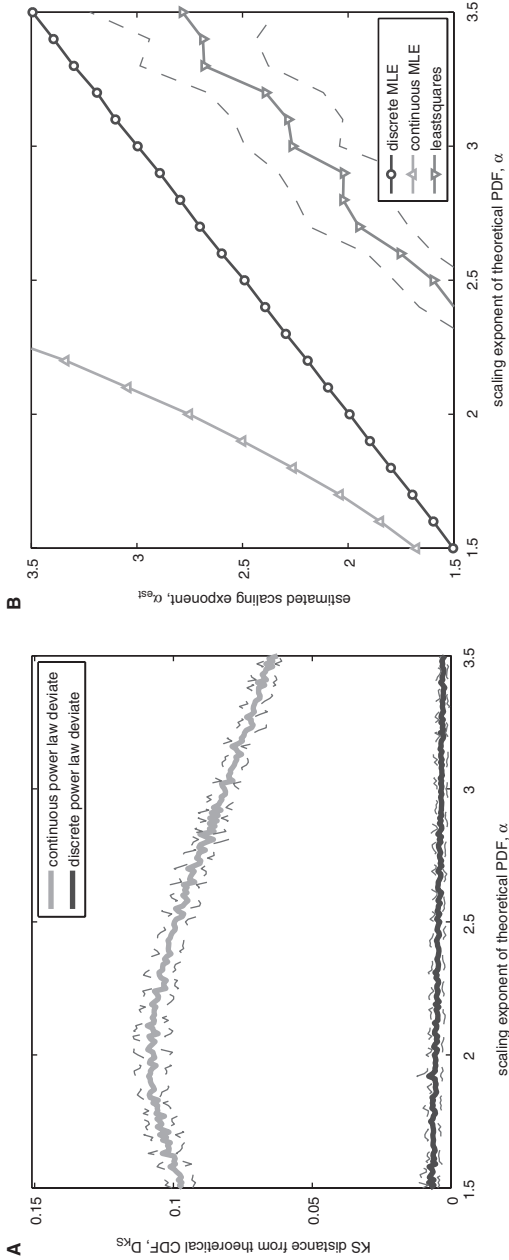## Generating Power-Law Distributed Data

Statistical modeling often necessitates the generation of random deviates with a specified distribution, for example, in simple null models or statistical hypothesis tests. Newman (2005) gives a simple analytic formula, derived using the transformation method in Press et al. (1992), for converting a uniform deviate into a continuous power-law deviate:

$$x = x_{min}(1 - r)^{-1/(\alpha-1)}$$

where x is distributed as a real number over the interval $[x_{\min}, \infty)$, and r is a uniform deviate. Although it may be tempting to simply take the integer portion of each deviate x to obtain a discrete power law, the resulting distribution will actually differ quite strongly from what is desired: such a procedure shifts a significant amount of probability mass from smaller to larger values, relative to the corresponding theoretical discrete power-law distributed deviate.

A more satisfying approach is to use a deviate generator specifically designed for a discrete power law. Because the discrete form does not admit a closed-form analytical solution via the transformation method like the continuous form, the generator must instead take an algorithmic approach to convert uniform deviates via the inverse cumulative density function of the discrete power law. Such an approach is a standard practice, and fast algorithms exist for doing so (Press et al. 1992). To illustrate the differences between these two power-law deviate generators, we show in Figure 6a that the latter approach produces distributions that are significantly closer to the desired theoretical one than does the former method, and it is the latter that we use for our statistical studies in the main text.

Figure 6
Comparison of Techniques for Generating and Fitting Power-Law Distributions

Note: (a) The closeness in the sense of the Kolmogorov-Smirnov goodness-of-fit measure of power-law distributed deviates generated using the two methods described in the text to the target distribution, a discrete power law with scaling parameter $\alpha = 2.5$. Results are for $x_{min} = 1$ and $n = 10,000$, with similar results holding for other values, although the difference decreases as $x_{min} \to \infty$. Quite dramatically, the discrete deviate generator does a significantly better job at matching the theoretical distribution than does the continuous method discussed in the text. (b) The results of using the three methods discussed in the text for estimating the scaling parameter of discrete power-law distributed data, with parameters $x_{min} = 1$ and $n = 10,000$; similar results hold for other values, although the estimates get increasingly noisy as the number of observations shrinks, and the two maximum likelihood estimators increasingly agree as $x_{min} \to \infty$. Error bars are omitted when they are less than the size of the series symbol.

Estimating Scaling Parameters from Data

Since Richardson (1948) first considered the scale invariance in the frequency and severity of wars, statistical methods for characterizing power laws have advanced significantly. The signature feature of a tail distribution that decays as a power law is a straight line with slope α on doubly logarithmic axes. As such, a popular method of measuring the scaling exponent α has been by a least-squares regression on log-transformed data, that is, one takes the log of both the dependent and independent variables, or one could bin the data into decades, and then measures the slope using a least-squares linear fit. Unfortunately, this procedure yields a biased estimate for the scaling exponent (Goldstein, Morris, and Yen 2004). For continuous power-law data, Newman (2005) gives an unbiased estimator based on the method of maximum likelihood; however, it too yields a biased estimate when applied to discrete data such as ours. Goldstein, Morris, and Yen (2004) studied the bias of some estimators for power-law distributed data, and, also using the method of maximum likelihood, give a transcendental equation whose solution is an unbiased estimator for discrete data. In our main study, we use a generalization of this equation as our discrete maximum likelihood estimator.

To give the reader a sense of the performance of these methods, we show in Figure 6b the results of applying them to simulated data derived from the discrete generator described above. Quite clearly, the discrete maximum likelihood estimator yields highly accurate results, with the other techniques either over- or underestimating the true scaling parameter, sometimes dramatically so. Johnson et al. (2006) have also studied the accuracy of these estimators but apparently only for data derived from the continuous deviate generator described above.

The discrete maximum likelihood estimator of Goldstein, Morris, and Yen (2004) assumes that the tail encompasses the entire distribution. A generalization of their formula to distributions where the tail begins at some minimum value $x_{\min} \geq 1$ follows, and the value of α_{ML} that satisfies this equation is the discrete maximum likelihood estimator:

$$\frac{\zeta'(\alpha, x_{\min})}{\zeta(\alpha, x_{\min})} = -\frac{1}{n} \sum_{i=1}^{n} \log x_I,$$

where the x_i are the data in the tail, n is the number of such observations, and $\zeta(\alpha, x_{min})$ is the incomplete Riemann zeta function. If desired, the latter can be rewritten as $\zeta(\alpha) - H^{\alpha}_{x_{\min}}$, being the difference between a zeta function and the $x_{\min}$th harmonic number of order α. When $x_{\min} = 1$, the left-hand side reduces to $\zeta'(\alpha)/\zeta(\alpha)$, the values of which can be calculated using most standard mathematical software. Alternatively, one can numerically maximize the log-likelihood function itself,

$$L(\alpha \mid x) = -n \log \zeta(\alpha, x_{\min}) - \alpha \sum_{i=1}^{n} \log x_i,$$

which may be significantly more convenient than dealing with the derivative of the incomplete zeta function. This approach is what was used in both the present study and in the preliminary Clauset and Young (2005) study of terrorism.

These equations assume that the range of the scaling behavior, that is, the lower bound $x_{\min}$, is known. In real-world situations, this value is often estimated visually, and a conservative estimate of such can be sufficient when the data span a half-dozen or so orders of magnitude. However, the data for many social or complex systems only span a few orders of magnitude at most, and an underpopulated tail would provide our tools with little statistical power. Thus, we use a numerical method for selecting the $x_{\min}$ that yields the best power-law model for the data. Specifically, for each $x_{\min}$ over some reasonable range, we first estimate the scaling parameter α_{ML} over the data $x \geq x_{\min}$ and then compute the Kolmogorov-Smirnov (KS) goodness-of-fit statistic between the data being fit and a theoretical power-law distribution with parameters α_{ML} and $x_{\min}$. We then select the $x_{\min}$ that yields the best such fit to our data. For simulated data with similar characteristics to the MIPT (2006) data, we find that this method correctly estimates both the lower bound on the scaling and the scaling exponent. Mathematically, we take

$$x_{\min} = \min_{y} \left[\max_{x} |F(x; \alpha_{\mathrm{ML}}, y) - \hat{F}(x; y)| \right],$$

where $F(x; y, \alpha_{ML})$ is the theoretical cumulative distribution function (cdf) for a power law with parameters α_{ML} and $x_{\min} = y$, and $\hat{F}(x; y)$ is the empirical distribution function (edf) over the data points with value at least y. In cases where two values of y yield roughly equally good fits to the data, we report the one with greater statistical significance.

Once these parameters have been estimated, we first calculate the standard error in α via bootstrap resampling. The errors reported in Tables 1 and 2, for instance, are derived in this manner. Finally, we calculate the statistical significance of this fit by a Monte Carlo simulation of n data points drawn a large number of times (e.g., at least 1,000 draws) from $F(x; \alpha_{ML}, x_{min})$, where α_{ML} and $x_{\min}$ have been estimated as above under the one-sided KS test. Tabulating the results of the simulation yields an appropriate table of p-values for the fit, by which the relative rank of the observed KS statistic can be interpreted in the standard way.

As mentioned in the text, there are many heavy-tailed distributions, for example, the q-exponential $e_q^{-\alpha x}$, the stretched exponential $e^{-\alpha x^{\beta}}$, the log-normal, and even a different two-parameter power law $(c + x)^{-\alpha}$. For data that span only a few orders of magnitude, the behavior of these functions can be statistically indistinguishable; that is, it can be hard to show that data generated from an alternative distribution would not yield just as good a fit to the power-law model. As such, we cannot rule out all Type II statistical errors for our power- law models. On the other hand, we note that for the distributions described in "Frequency-Severity Distributions for Attacks since 1968," the statistical power test versus a log-normal model indicates

that the power law better represents the empirical data. In some sense, the particular kind of asymptotic scaling in the data is less significant than the robustness of the heavy tail under a variety of forms of analysis. Simply the fact that the patterns in the real-world severity data deviate so strongly from our expectations via traditional models of terrorism illustrates that there is much left to understand about this phenomenon, and our models need to be extended to account for the robust empirical patterns we observe in our study.

Notes

1. See figures for the Financial Times Stock Exchange 100 index of the 100 largest companies listed on the London Stock Exchange at http://www.econstats.com/eqty/eq_d_mi_5.htm.

2. A straight line on doubly logarithmic axes is a necessary but not sufficient condition for a distribution to be a power law; for example, when we have only a small number of observations from an exponentially distributed variable, it can appear roughly straight on doubly logarithmic axes.

3. The ccdf is preferable to the probability distribution function (pdf), since the latter is significantly noisier in the upper tail, exactly where subtle variations in behavior can be concealed. If a distribution scales, it will continue to do so on the ccdf.

4. The Kolmogorov-Smirnov test evaluates whether observed data seem a plausible random sample from a given probability distribution by comparing the maximum difference between the observed and the expected cumulative distributions.

5. In 1998, the management of the database was transferred from the RAND Corporation to the National Memorial Institute for the Prevention of Terrorism (MIPT), which resulted in several observable differences in the database records. For instance, although some purely domestic events appear prior to 1998, such as the 1995 Oklahoma City bombing, domestic events make up a significant fraction of the events entered subsequent to 1998, suggesting that the true number of events for some period directly prior to 1998 is greater than we observe in the database. Although this effect could create problems for analyses that count incidents in a simple way, it does not affect the scale invariant shape of the frequency-severity distribution, primarily, we believe, because the large events that comprise the tail of the distribution were the least susceptible to any underreporting bias. We shall explore this point more in the next section.

6. The attacks of September 11, 2001, are considered unconventional.

7. Critical phenomena characterize a phase transition such as the evaporation of water, while an equilibrium critical phenomenon is one in which the critical state is a global attractor of system dynamics.

8. A trivial spatial model for the frequency-severity scale invariance would be a tight connection with size of the targeted city and the number of casualties. That is, as we saw earlier, large city populations are distributed as a power law, and we might suppose that an event's severity is proportional to the size of the target city. If target cities are chosen roughly uniformly at random, an obviously unrealistic idealization, then a power law in the frequency-severity statistics follows naturally. Tabulating population estimates for cities in our database from publicly available census data, we find that the correlation between an event's severity and the target city population is very weak, $r = 0.2(2)$, for deaths and $r = 0.2(1)$ for total severity, where the number in parentheses is the standard error from a bootstrap resampling of the correlation calculation.

9. On the planning for the September 11, 2001, attacks, see the summary of a documentary aired by Al Jazeera at http://archives.cnn.com/2002/WORLD/meast/09/12/ alqaeda.911.claim/.

References

Arce, M., Daniel G., and Todd Sandler. 2005. Counterterrorism: A game-theoretic analysis. *Journal of Conflict Resolution* 49 (1): 138-200.

Bak, Per, and Chao Tang. 1989. Earthquakes as a self-organized critical phenomena. *Journal of Geophysical Research* 94 (B11): 15635-7.

Bak, Per, Chao Tang, and Kurt Wiesenfeld. 1987. Self-organized criticality: An explanation of 1/f noise. *Physical Review Letters* 59 (4): 381-4

Biggs, Michael. 2005. Strikes as forest fires: Chicago and Paris in the late 19th century. *American Journal of Sociology* 111 (1): 1684-714.

Bogen, Kenneth T., and Edwin D. Jones. 2006. Risks of mortality and morbidity from worldwide terrorism: 1968-2004. *Risk Analysis* 26 (1): 45-59.

Cederman, Lars-Erik. 2003. Modeling the size of wars: From billiard balls to sandpiles. *American Political Science Review* 97 (1): 135-50.

Cioffi-Revilla, Claudio. 1991. On the likely magnitude, extent, and duration of the Iraq-UN war. *Journal of Conflict Resolution* 35 (3): 387-411.

Clauset, Aaron, and Maxwell Young. 2005. Scale invariance in global terrorism. Preprint. *http://arxiv.org/abs/physics/0502014* (accessed on February 3, 2005).

Congleton, Roger D. 2002. Terrorism, interest-group politics, and public policy. *Independent Review* 7 (1): 47-67.

Crook, Clive. 2006. The height of inequality. *The Atlantic Monthly* 298 (2): 36-37.

Enders, Walter E., and Todd Sandler. 2006. *The political economy of terrorism.* Cambridge, UK: Cambridge University Press.

Farmer, J. Doyne, and John Geanakoplos. 2006. Power laws in economics and elsewhere. Unpublished manuscript. Santa Fe Institute, NM/Yale University, New Haven CT.

Federal Bureau of Investigation, Counterterrorism Threat Assessment and Warning Unit, Counterterrorism Division. 1999. Terrorism in the United States, 1999. US Department of Justice, Washington, DC. http://www.fbi.gov/publications/terror/terror99.pdf (accessed on January 22, 2005).

Gartner, Scott S. 2004. Making the international local: The terrorist attack on the USS Cole, local casualties, and media coverage. *Political Communication* 21 (1): 139-59.

Goldstein, Michel L., Steven A. Morris, and Gary G. Yen. 2004. Problems with fitting to the power-law distribution. *European Physical Journal B: Condensed Matter and Complex Systems* 41 (2): 255-8.

Hoffman, Bruce. 1999. Terrorism trends and prospects. In *Countering the new terrorism*, edited by Ian O. Lesser, Bruce Hoffman, John Arquilla, David Ronfeldt, Michele Zanini, and Brian Michael Jenkins, 7-38. Santa Monica, CA: RAND.

Johnson, Neil F., Michael Spagat, Jorge Restrepo, Juan Bohorquez, Nicholás Suarez, Elvira Restrepo, and Roberto Zarama. 2005. From old wars to new wars and global terrorism. Preprint. *http://arxiv.org/abs/physics/0506213* (accessed on June 29, 2005).

Johnson, Neil F., Michael Spagat, Jorge A. Restrepo, Oscar Becerra, Juan C. Bohorquez, Nicholás Suarez, Elvira M. Restrepo, and Roberto Zarama. 2006. Universal patterns underlying ongoing wars and terrorism. Preprint. *http://arxiv.org/abs/physics/0605035* (accessed on May 3, 2006).

Kleiber, Christian, and Samuel Kotz. 2003. *Statistical size distributions in economics and actuarial sciences.* Hoboken, NJ: John Wiley.

Lacina, Bethany A. 2006. Explaining the severity of civil wars. *Journal of Conflict Resolution* 50 (2): 276-89.

Li, Lun, David Alderson, John C. Doyle, and Walter Willinger. 2005. Towards a theory of scale-free graphs: Definition, properties, and implications. *Internet Mathematics* 2 (4): 431-523.

Li, Quan. 2005. Does democracy promote or reduce transnational terrorist incidents? *Journal of Conflict Resolution* 49 (2): 278-97.

Malamud, Bruce D., Gleb Morein, and Donald L. Turcotte. 1998. Forest fires: An example of self-organized critical behavior. *Science* 281:1840-1.

Mickolus, Edward F., Todd Sandler, Jean M. Murdock, and Peter A. Flemming. 2004. *International terrorism: Attributes of terrorist events, 1968-2003 (ITERATE 3-4).* Dunn Loring, VA: Vinyard Software.

Mitzenmacher, Michael. 2004. A brief history of generative models for power law and lognormal distributions. *Internet Mathematics* 1 (2): 226-51.

National Memorial Institute for the Prevention of Terrorism. 2006. Terrorism knowledge base. http://www.tkb.org (access on June 18, 2006).

Navarro, Peter, and Aaron Spencer. 2001. September 11, 2001: Assessing the costs of terrorism. *Milken Institute Review* (Fourth Quarter): 16-31.

Newman, Mark. E. J. 2005. Power laws, Pareto distributions and Zipf's law. *Contemporary Physics* 46 (5): 323-51.

Overgaard, Per B. 1994. The scale of terrorist attacks as a signal of resources. *Journal of Conflict Resolution* 38 (3): 452-78.

Pape, Robert A. 2003. The strategic logic of suicide terrorism. *American Political Science Review* 97 (3): 343-61.

Press, William H., Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. 1992. *Numerical recipes in C: The art of scientific computing.* Cambridge, UK: Cambridge University Press.

Reed, William J., and Barry D. Hughes. 2002. From gene families and genera to income and Internet file sizes: Why power laws are so common in nature. *Physical Review E* 66 (6): 67103-7.

Reich, Walter. 1990. *Origins of terrorism.* Cambridge, UK: Cambridge University Press.

Richardson, Lewis F. 1948. Variation of the frequency of fatal quarrels with magnitude. *Journal of the American Statistical Association* 43 (244): 523-46.

Rosendorff, B. Peter, and Todd Sandler. 2005. The political economy of transnational terrorism. *Journal of Conflict Resolution* 49 (2): 171-82.

Sandler, Todd, and Daniel G. Arce M. 2003. Terrorism and game theory. *Simulation & Gaming* 34 (3): 319-37.

Sandler, Todd, and Harvey E. Lapan. 1988. The calculus of dissent: An analysis of terrorists' choice of targets. *Synthese* 76 (2): 245-61.

Serfling, Robert. 2002. Efficient and robust fitting of lognormal distributions. *North American Actuarial Journal* 6 (1): 95-109.

Shubik, Martin. 1997. Terrorism, technology, and the socioeconomics of death. *Comparative Strategy* 16 (4): 399-414.

Simon, Herbert. A. 1955. On a class of skew distribution functions. *Biometrika* 42 (3/4): 425-40.

United States Bureau of the Census. Census 2000. Washington, DC: Government Printing Office. http://www.census.gov/main/www/cen2000.html (accessed on September 18, 2006).

United States Department of State. 2004. Patterns of global terrorism, 2003. Washington, DC, U.S. Department of State, April. http://www.state.gov/documents/organization/31912.pdf (accessed on January 22, 2005).

Wilkinson, Paul. 1997. The media and terrorism: A reassessment. *Terrorism and Political Violence* 9 (2): 51-64.

Zipf, George K. 1949. *Human behavior and the principle of least effort.* Cambridge, MA: Addison-Wesley.