The Next Big Thing
Scale Invariance in Political Science

Arthur Spirling*

Abstract

'Power laws' suggest that events of a large magnitude will be rare, whilst small events will be much more common, and that a simple mathematical law relates 'severity' with frequency. We find that a wide variety of phenomena in political science are power law distributed. These empirical regularities are both unexpected and unexplained. More work on a general explanatory theory for these patterns is desirable.

Samuel Alito was recently confirmed to fill the vacant US Supreme Court position left by Sandra Day O'Connor, but should we be surprised that it took the President *two* attempts to get the Senate's assent to his choice? What kind of predictions can we give on the next democidal bloodletting by a despotic government? When, and to what extent, will governments respond to crises by launching undemocratic purges of their opposition? When alliances form in international relations—such as the "Coalition of the Willing" for Iraq—how large should we expect them to be?

At least two themes unite these seemingly disparate research topics. First, they refer to extremely complicated political science phenomena which are generally difficult to accurately predict. Second, they show remarkable 'scale invariance', a property that relates the *size* of an event to its *frequency*.[1]

In this paper, it is shown that many familiar political science systems of actor interaction are well described and characterized by a simple mathematical relationship, known as the power law. More formally, we can write that

$$\Pr(x) \propto x^{-\alpha} \tag{1}$$

where x is the size of the event, and α is a scaling exponent to be estimated.

This is a 'scale invariant' relationship in the sense that the above 'law' describing the distribution of x is always proportional to the same (invariant) α, regardless of the magnitude of the event in question. Heuristically, since α is a positive constant, it should be clear from (1) that we expect small sized events to occur regularly, but large events to be much rarer.

In the natural sciences, the literature on observations of scale invariance and related concepts like 'Zipf's law'[2] is, by now, enormous. Power law distributions appear in the

---

[1]Depending on the object or system under study, 'size' can be measured in different ways: we could quantify a war by the number of dead; an earthquake's severity is measured by its place on the Richter scale; the business of a barber's shop could be recorded by the volume of customers it received today.

[2]Zipf's law relates frequency to a *rank ordering* of event size (in the original case, word length in various

description of phenomena as diverse as condensed matter physics, earthquakes, volcanoes, forest fires, avalanches and network theory. Economists have studies the size of cities (Gabaix 1999) and fluctuations in the stock market (Gabaix, Gopikrishnan, Plerou & Stanley 2003) using power laws. In recent times, sociologists have used similar concepts to explore historical patterns of labor strikes (Biggs 2005).

In political science, such work is practically absent with the exception of the study of combat. In this latter case, Richardson (1948, 1960) considered international and domestic cases of 'deadly quarrels' (wars, murders and the like). Using a logarithmic scale with categories of $3, 4, 5, 6$ and 7, he classified the various events by their causality numbers. So, for example, the Second World War (1939–1945) with over a million casualties was placed (with the Great War, 1914–1918) into the '7' bin. He discovered a remarkable regularity: for every ten-fold increase in the 'size' of the event—e.g. a move from bin '3' to bin '4', or from bin '6' to bin '7'—the number of such events decreased by around 2.5 times relative to the previous category. This result has been remarkably stable even in the face of new data (Cederman 2003), and has been extended in recent times to the case of terrorism (Clauset & Young 2005, Johnson, Spagat, Restrepo, Bohorquez, Suarez, Restrepo & Zarama 2005) which also shows power law properties.

What is the point of this type of analysis? Three possibilities seem immediate: first, it *explores and summarizes* seemingly patternless data as a one parameter distribution. Second, *prediction*: even if an x of a particular size has never been observed (or only very rarely), knowledge of α allows researchers to predict with reasonable certainty $\Pr(x)$. Third, by imposing order on seemingly complex systems, these empirical regularities suggest new ways to *theorize* about our discipline and its mechanisms.

This paper demonstrates that power laws hold in many areas beyond Richardson's (1960) regularity. Data is drawn from the three primary fields of political science study: American

languages)

3

politics, Comparative politics and International Relations. In the next section, I present a statistical method for fitting power laws. Subsequently, I suggest that Senate acceptance or rejection of Supreme Court appointments has an interesting power law relationship. I then deal with the frequency and size of various government purges and democides, before turning attention to coalitions in international relations. In every case, though the results are 'messier' than those found by analysts working on larger data sets and in the natural sciences, evidence of power law behavior is demonstrated. Though the essence of this paper is empirical exploration, the final section concludes by pointing out some possible starting points for modeling the mechanisms at work. It is generally suggested that many further political science avenues of research are wide upon to such techniques as displayed here.

How to Determine Whether Data is Power Law Distributed

Suppose we have data on some political science 'event.' These events could be the number of committees to which a bill is referred by the US House Speaker, the number of 'rebels' in a House of Commons parliamentary roll call or the number of people killed by some terrorist act.

Assume for now that events are distributed as a random variable X and events of size x are drawn iid. Denote the support of X as $\mathrm{supp}(x)$ and assume it is discrete,

$$\mathrm{supp}(x) = \{1, 2, \ldots, n\}$$

where $n = x_{[n]}$, i.e., the most severe ('largest') observation in the data set. Notice that events of size '0' are not admissible, since they are logical impossibilities: e.g. the Speaker must refer a bill to *at least one* committee for the referral event to have taken place.

If, as in this case, $x_{[1]} = 1$, then a power law distribution would imply

$$\Pr(X = x | \alpha) = \frac{x^{-\alpha}}{\zeta(\alpha)} \tag{2}$$

where $\zeta(\alpha)$ is the Riemann zeta function, $\sum_{k=1}^{\infty} k^{-\alpha}$. Of course, summing to ∞ is somewhat impractical so the zeta function is approximated as $\sum_{k=1}^{100000} k^{-\alpha}$.

The likelihood is then

$$\mathcal{L}(\alpha) = \frac{\prod_{i=1}^{n} x_i^{-\alpha}}{\zeta(\alpha)^n} \tag{3}$$

and the log-likelihood is

$$\ell(\alpha) = -\alpha \left(\sum_{i=1}^{n} \ln x_i \right) - n \ln \left(\sum_{k=1}^{\infty} k^{-\alpha} \right). \tag{4}$$

The method proceeds as follows: obtain a bootstrapped maximum likelihood estimate of α, $\hat{\alpha}$. Use a one-sample Kolmogorov-Smirnov test to obtain a p-value for the $\hat{\alpha}$ under consideration. That is, use a KS test to compare the actual distribution of data to one that is power law distributed with (single) parameter $\hat{\alpha}$. Readers may be more familiar with comparisons via some variant of a χ^2 test, but such an approach is not feasible here because of a very low number (less than five) of observations in certain categories.

In the cases presented below, it is *not* possible to reject the null hypothesis that the respective events are drawn from a power law distribution.[3] It is also the case that a power law distribution is a better fit than a zero-truncated Poisson for every example. In general a truncated Poisson simply does not predict the very rare 'large' events.

---

[3]Notice that this estimation procedure differs for that utilized by Clauset & Young (2005) since there is no estimation of $x_{\min}$, the smallest event size that can be included in the power law distribution.

Judge Ye Not

At the time of writing, Samuel Alito had just been confirmed to the vacant position of Justice of the Supreme Court of the United States, following Sandra Day O'Connor's July 1, 2005 announcement that she planned to retire from duty. Article II (Section 2) of the Constitution mandates that Supreme Court Justices are nominated by the President of the United States with the "advice and consent" of the Senate. Quite what this consultation process entails is open to speculation, and has been interpreted in widely different ways. In the event, the Senate may either confirm or reject any nomination.

Generally, the Senate approves the President's nominee for a vacant spot. That is, the President only needs to go through the whole process of nominating someone for a vacant spot *once per vacancy*. But sometimes, life is more interesting. For example, in June 1987, Lewis Powell, a relative liberal, resigned from the Supreme Court. His proposed replacement, Robert Bork, a conservative, inspired some of the most heated and impassioned debate in Supreme Court history. Within an hour of the nomination, Senator Edward Kennedy condemned President's Reagan's choice in the strongest possible terms on the floor of the legislature and Bork's Senate nomination hearing lasted 5 days, the longest ever. He was eventually rejected 58–42 by the Senate. The Bork story itself pails into comparison with the attempted replacement of Smith Thompson. Resigning in 1843, some 6 replacements were proposed before a successor—Samuel Nelson—was appointed.

The data itself is available from the website of the US Senate. There are some coding issues: in particular, what counts as a failed nomination attempt. Sometimes, as with Bork, the Senate rejects the nominee, but often, the nomination simply lapses because a President finishes his tenure without a Senate decision. Similarly, rather than face a (likely rejection) vote on his chosen replacement, a President may withdraw a nominee. In rare circumstances, a candidate may actually decline the offer of a Justice position, though this has not occurred

since 1837 when a new seat became available through creation, rather than retirement. All these cases are treated in the same way as a flat Senate rejection: that is, they count as 'attempts to replace a Justice'.

For the Justice nominee data, the bootstrapped $\hat{\alpha}$ is 2.96, and the p-value is 0.58. This is good news: it is not reasonable to reject the null of a power law distribution, and there is strong evidence that this distribution is a good approximation to the actual data. To visualize this fit, Figure 1 gives the actual (solid lines, triangle plot characters) values and power law predicted (broken line, filled circles) values for the data.

It should be evident from Figure 1 that the scale invariant model with parameter $\alpha = 2.96$ performs well for this data. A zero-truncated Poisson was also fitted to the data via maximum likelihood estimation ($\hat{\lambda} = 0.66$), and does poorly relative to a power law, especially in predicting the largest 'sizes.' The zero-truncated Poisson fit is represented by the broken line and unfilled square plotting characters.

It may not be immediately clear from the graphic, but the truncated Poisson does particularly badly at predicting the larger sizes. Table 1 compares the proportions predicted by the power law model and the zero truncated Poisson to the actual proportions. For example, $\Pr(x = 6) = 0.0092$ in the actual data, whilst the power law fit predicts 0.0071, and the truncated Poisson predicts 0.0011.

x	actual proportion	power law pred.	trunc. Poisson pred.
1	0.7890	0.8264	0.7060
2	0.1376	0.1062	0.2330
3	0.0183	0.0320	0.0513
4	0.0367	0.0136	0.0085
5	0.0092	0.0071	0.0011
6	0.0092	0.0041	0.0001

Table 1: SCOTUS: Actual and predicted proportions, by value of x
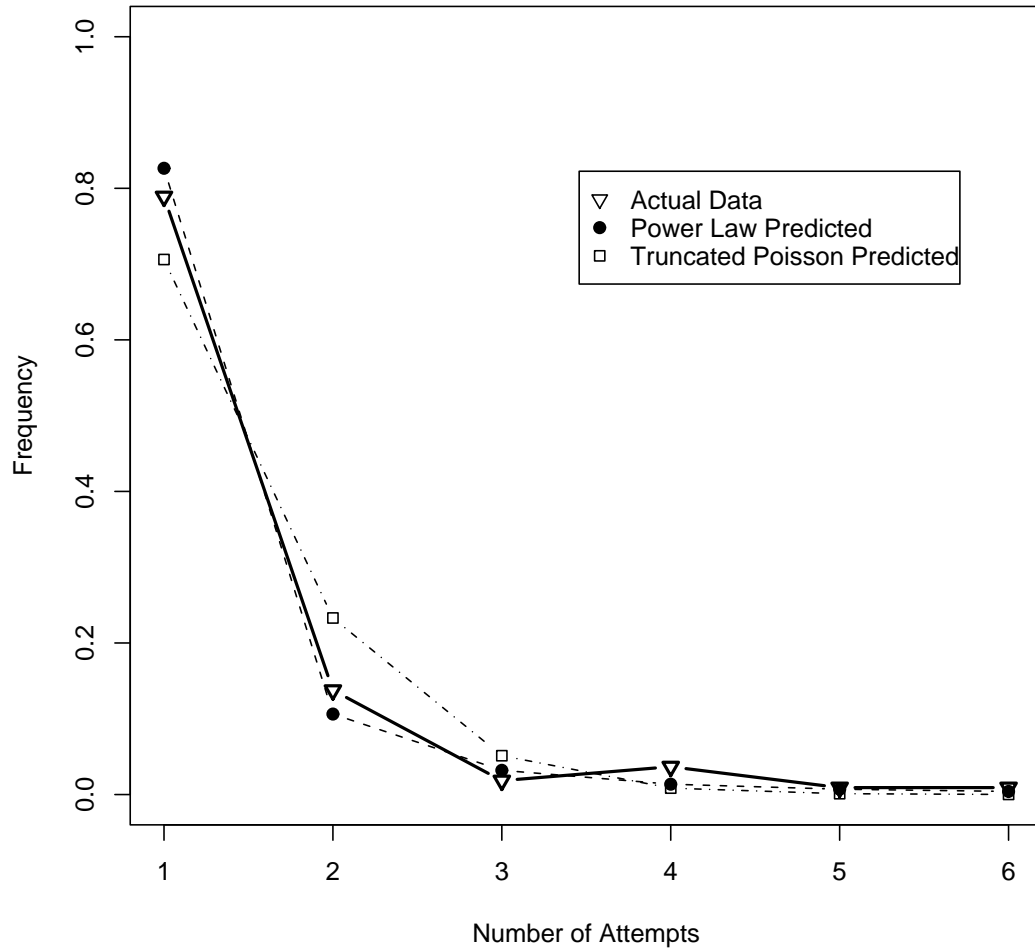
Figure 1: Actual and fitted densities for Supreme Court Justice replacement nomination attempts, 1790–2005

The Urge to Purge

Around 170 million people were murdered by their own governments in the 20th century. This is several times the number that were killed in war over a similar period. The USSR and the People's Republic of China particularly excelled in the morbid competition to slaughter their own citizens, clocking up an obscene 61 and 35 million respectively. Nazi Germany 'only' managed 20 million by comparison.[4]

Despotic democides—the mass murder of citizens by their governments—are sadly not uncommon. R. J. Rummel has collated an extensive database of these horrific events and his data was used to compile figures 2 and 3.

Figures 2 and 3 compare the log of the 'size' (number of people killed) with the log of the rank of the observation. Hence, the USSR is ranked as '1', the PR China as '2', and so on. The straight lines are simply linear regression slopes. The fits of these lines are remarkable: R^2 of 0.94 and 0.97 respectively.

Readers may mistakenly assume that this is simply a mathematical quirk of the size-rank relationship. These same readers should disabuse themselves by casting an eye over Figure 4 and Figure 5, which respectively show log–log plots for a 1000 uniform and a 1000 Poisson ($\lambda = 10$) distributed observations. Evidently, the pattern is far from a neat line, and the R^2 are a commensurately poor 0.44 and 0.67 respectively.
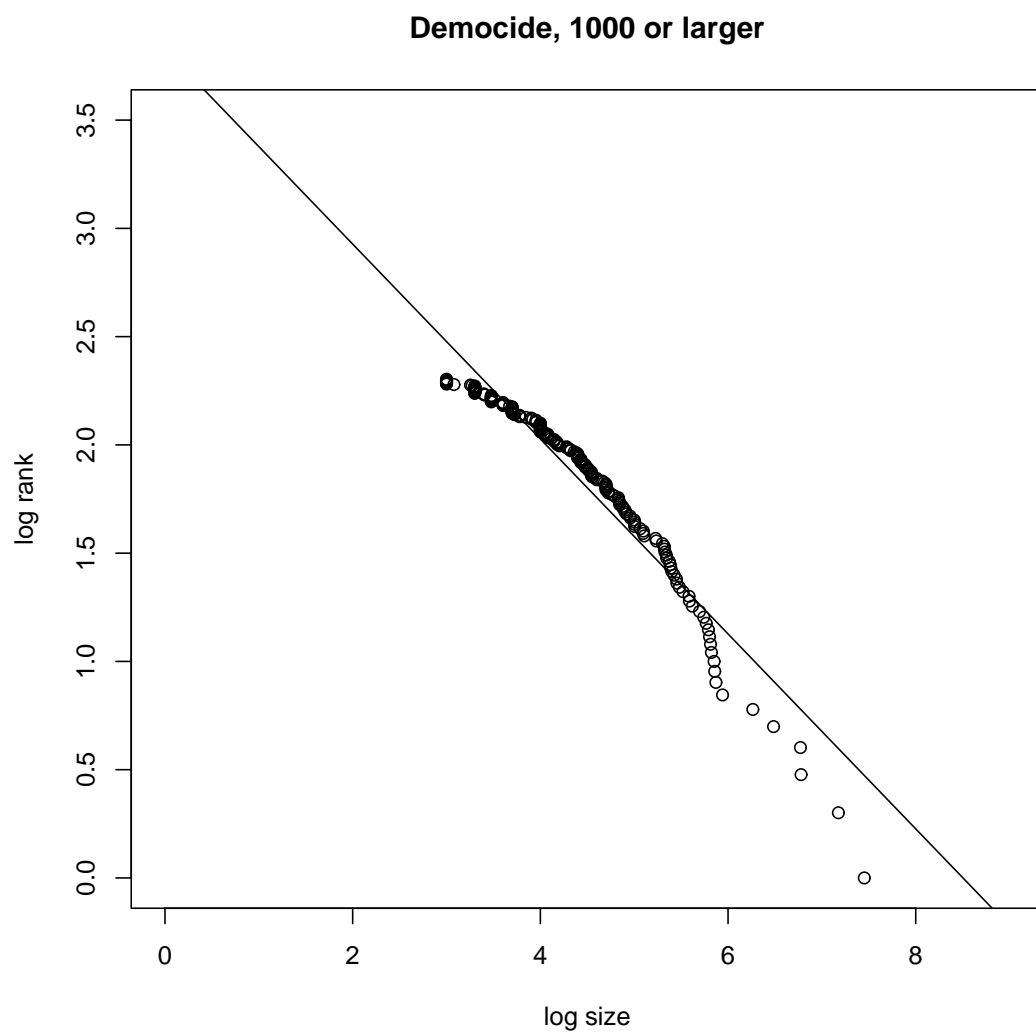
---

[4] `http://www.hawaii.edu/powerkills/20TH.HTM`

Democide, 1000 or larger

Figure 2: Zipf plot for Democide data ($\geq$ 1000 victims), 1900–1987

Democide, 10000 or larger

Figure 3: Zipf plot for Democide data ($\geq$ 10000 victims), 1900–1987
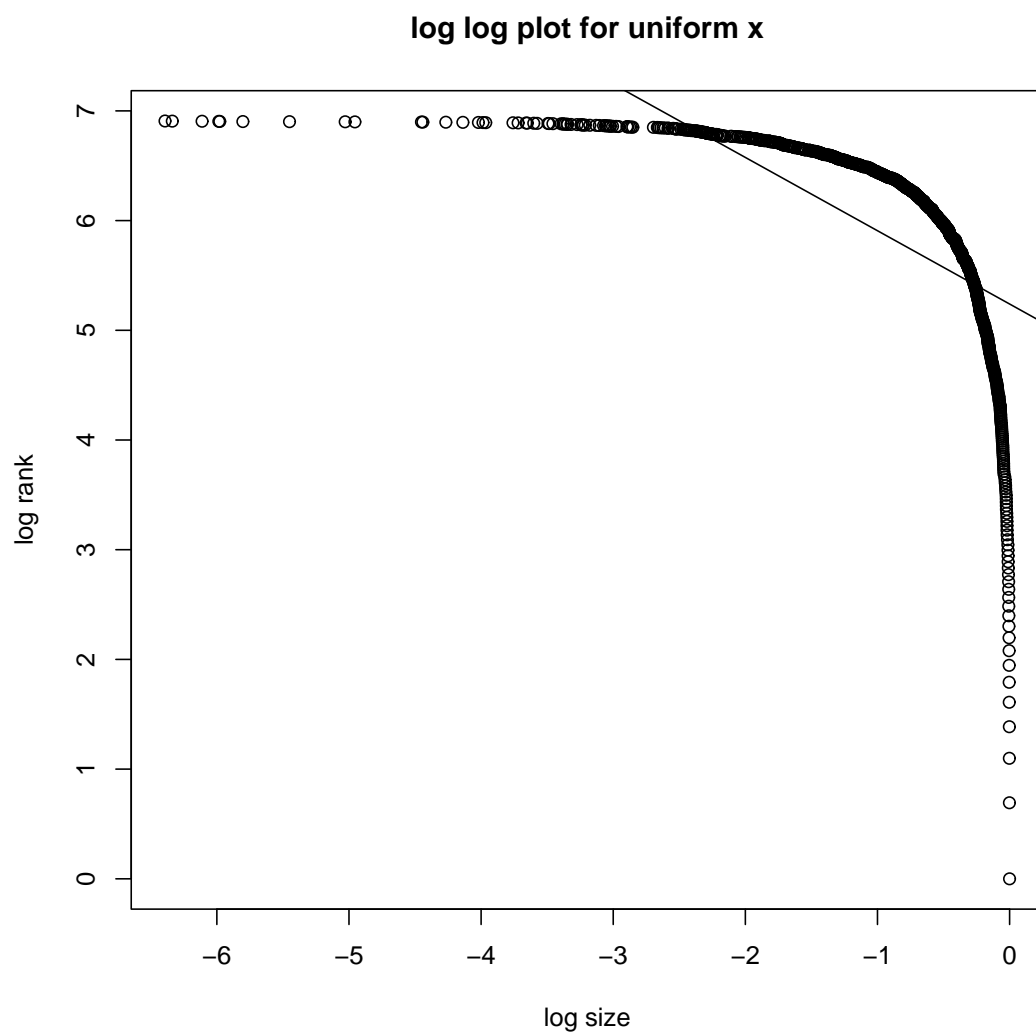
11

log log plot for uniform x

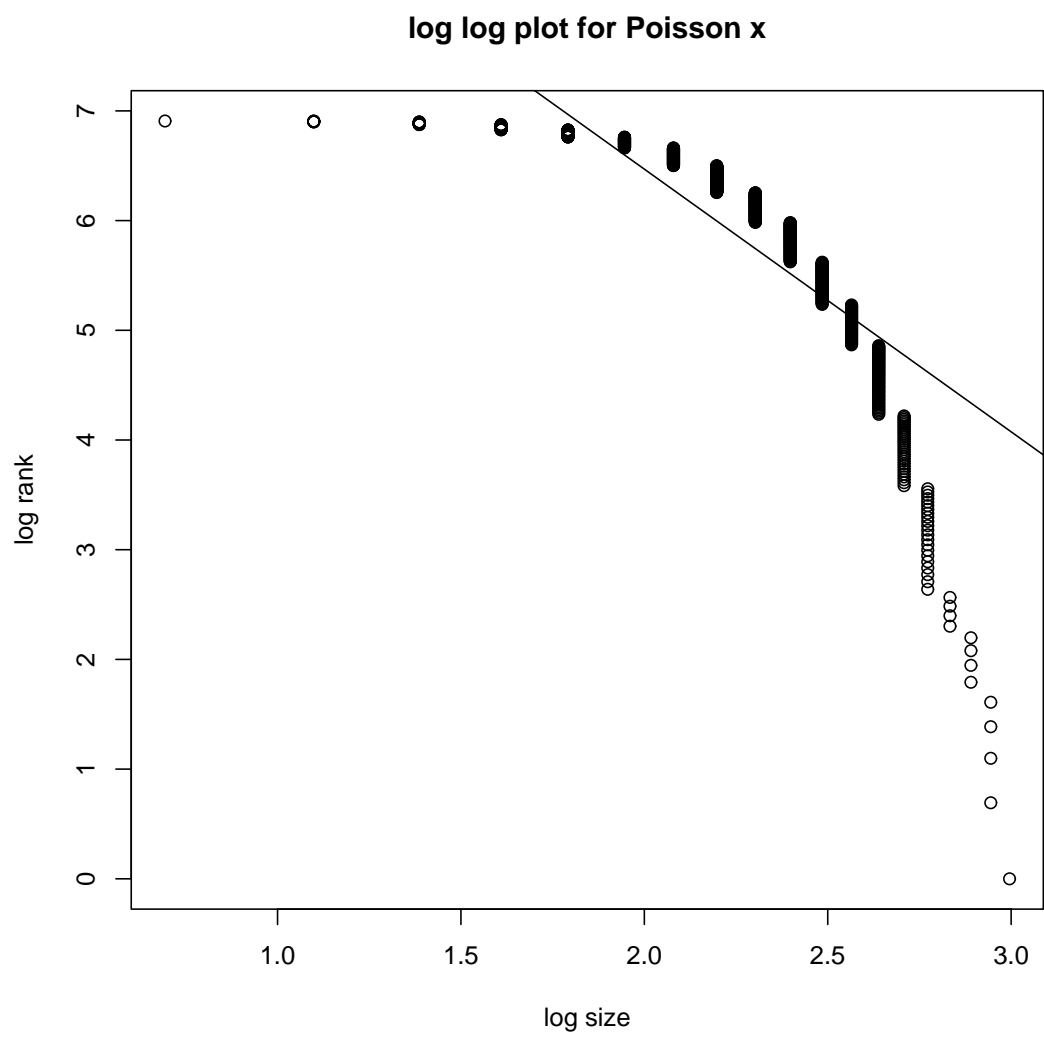Figure 4: Zipf plot for uniform data

12

Figure 5: Zipf plot for Poisson data

Technically speaking, though these are 'Zipf' plots, the relationships shown for democide do not obey Zipf's *law*, because the slope of the line is *not* -1. Nonetheless, the plots *are* indicative of some scale invariance or power law behavior in general (Mandelbrot 1997).

Unfortunately, the methodology described above does not facilitate the precise calculation of a power law exponent for these data, since the support of x does not begin at 1. As a substitute of sorts, scholars who visit the website of the ICPSR, can readily download data pertaining to two studies (the data of which have been merged) by Rummel and Tanter that do include some variables that might be considered comparable to the democide counts above.

The data deals with domestic and—non-war—IR conflict for a total of 77 countries for the years 1955–1960. These include the frequency of government led purges of the opposition or other bodies. Importantly, this variable shows power law distribution behavior.

A word about the treatment of the data is in order before the results are presented. The study, unfortunately, does not record the number of individuals involved in an event,[5] only the frequency *per country year* of different numbers of such events. Hence, for the purges variable, the size varies from one "systematic elimination by political elites. . . of opposition" per country year, to a maximum of five, for 1955 through 1960. For this variable, α is estimated as 2.83, with a *p*-value of 0.55. Figure 6 gives a graphical representation of the fit, with the solid lines as the actual data, and the filled circles as the power law predictions.

The fit here is not quite as impressive as that for, say, the Supreme Court nominations procedure, but it is decent. This is surprising for two reasons: the first is that the sample size is very small ($n = 41$), the second is that the measurement of 'size' of the event is very crude. That any power law-like relationship is found in such a circumstance is remarkable. The zero-truncated Poisson λ for this example is estimated as 0.49. Again, the fit is worse

---

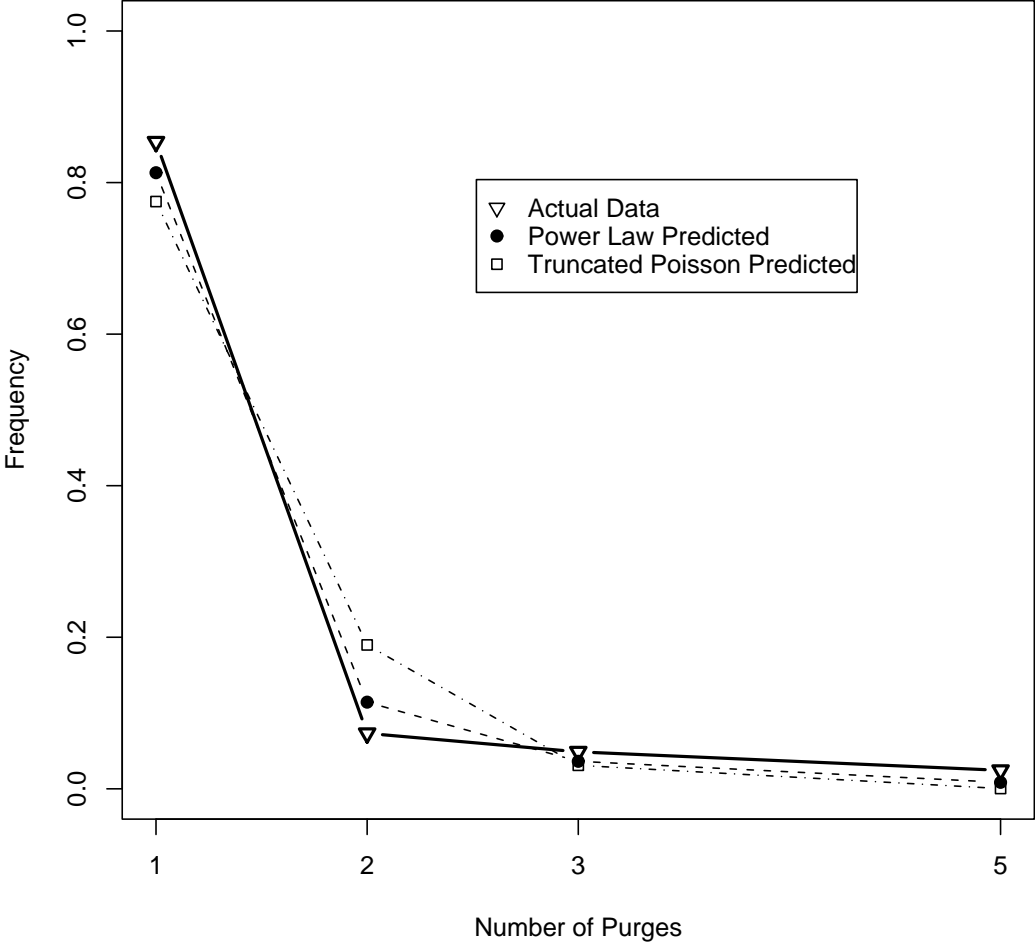[5]Other than above some requisite threshold for inclusion in the study.

Figure 6: Power law fit for purges data

that for the power law, and can be seen as the open square characters. Table 2 allows a direct comparison of the predicted probabilities for each value of x.

x	actual proportion	power law pred.	trunc. Poisson pred.
1	0.8537	0.8130	0.7749
2	0.0732	0.1143	0.1899
3	0.0488	0.0363	0.0310
5	0.0244	0.0086	0.0004

Table 2: Purges: Actual and predicted proportions, by value of x

Getting it together: Coalitions

International relations theorists commonly see coalitions, or alliances, in two circumstances. First, when countries go to war, and second, when countries want to act together in peace time. This latter case might apply in cases of sanctions or cross-national efforts to fight natural disasters.

Presumably the size of the coalition that forms in a particular circumstance is determined by a myriad of factors: the importance of trade relations between states; common ideological bonds; family relations between monarchies; international organization mandates and so on.

There are two immediately available sources of data for investigating the probability of a power law distribution for coalitions in IR. First, the Correlates of War project has data on alliances 1816–2000 (Gliber & Sarkees 2004), from which the number of states in each alliance is easily calculated. Second, Martin (1993) has recorded sanction data 1939–1983.

Before launching into the empirical investigation, readers should note that the minimum alliance size *by definition* is two. To ensure that the data conditions for equation (2) hold, the size of alliances is conceptually adjusted to the number of partners each member of the alliance had. Hence, a two nation alliance is a situation where each country has one partner; a five nation alliance sees each member state with four partners and so on.

16

Alliances come in several varieties according to "commitment type" which is captured in the COW coding. First, there are defence pacts; second are neutrality pacts; finally there are nonaggression pacts. In general we might imagine an ordinal ranking of commitment, but the COW codebook makes clear that this need to be the case; for the purposes of this paper, we only consider type 1.

The estimated α here is 2.32 and the p-value for the for the first alliance type is 0.79. Figure 7 gives the predicted and actual density of data in with the same line coding as previously. The zero-truncated Poisson in this case had a λ estimated as 1.85, and fails to predict the large values for the numbers of partners. Though this can be seen from the graphic—recall that the open squares are the truncated Poisson fit—it is even more obvious from Table 3: notice that the truncated Poisson fails to place positive probability on any number of partners over eight.

x	actual proportion	power law pred.	trunc. Poisson pred.
1	0.7128	0.7037	0.3452
2	0.1170	0.1409	0.3193
3	0.0745	0.0550	0.1969
4	0.0266	0.0282	0.0911
5	0.0213	0.0168	0.0337
7	0.0053	0.0077	0.0027
8	0.0053	0.0057	0.0006
13	0.0053	0.0018	0.0000
15	0.0053	0.0013	0.0000
19	0.0160	0.0008	0.0000
20	0.0053	0.0007	0.0000
34	0.0053	0.0002	0.0000

Table 3: Allies: Actual and predicted proportions, by value of x

For the Martin data, for those cases where the US took the lead as 'sender' (approximately 65% of the cases), $\hat{\alpha}$ is 2.00 and the p-value is 0.75.[6] Figure 8 gives the predicted and actual density of data for these cases with the same line coding as previously.

---

[6]For the non-US led cases, the pattern is not power law distributed
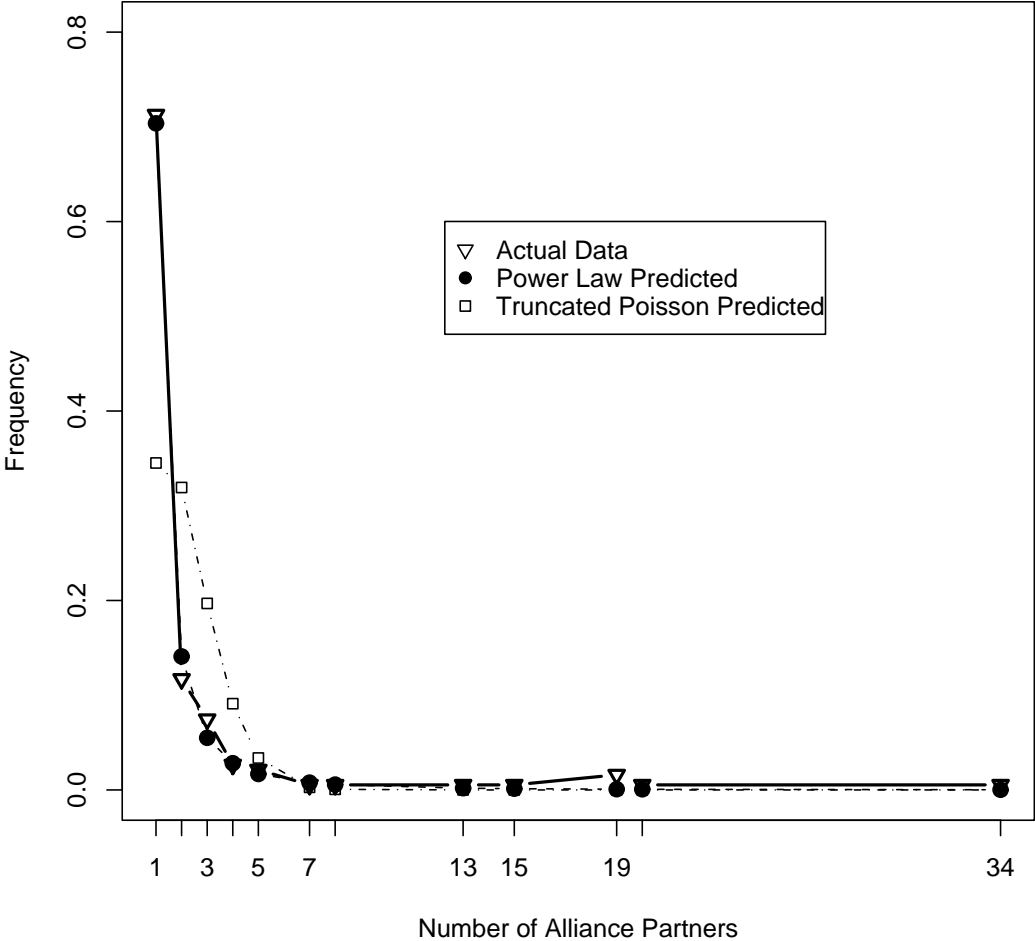
Alliances

Figure 7: Actual and Fitted Densities for Type 1 Alliances (number of partners, 1816–2000)
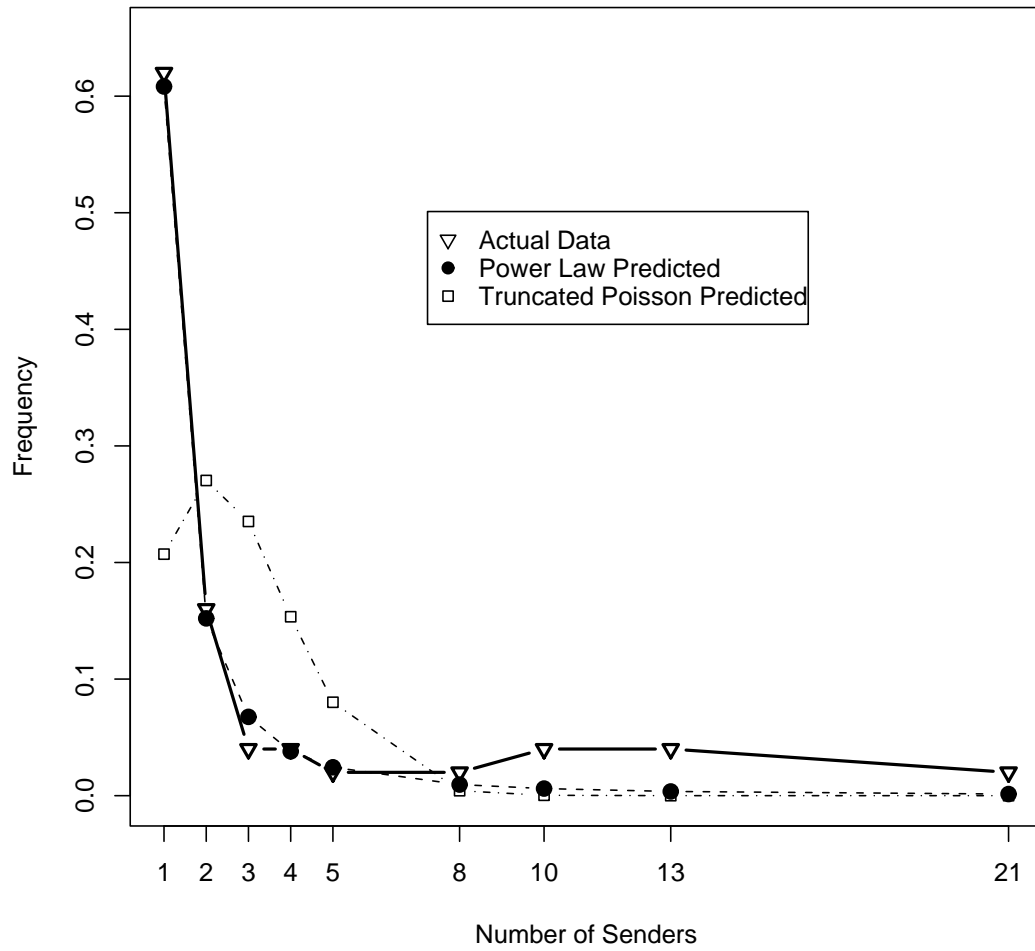
Sanction Senders

Figure 8: Actual and Fitted Densities for US led Sanction coalitions (number of senders, 1939–1983)

The zero-truncated Poisson was a very poor fit, with an estimated $\lambda = 2.61$. Once again, the truncated Poisson does poorly at fitting the larger 'sizes', as Table 4 makes clear.

x	actual proportion	power law pred.	trunc. Poisson pred.
1	0.62	0.6083	0.2072
2	0.16	0.1521	0.2703
3	0.04	0.0676	0.2352
4	0.04	0.0380	0.1535
5	0.02	0.0243	0.0801
8	0.02	0.0095	0.0042
10	0.04	0.0061	0.0003
13	0.04	0.0036	0.0000
21	0.02	0.0014	0.0000

Table 4: Sanction senders: Actual and predicted proportions, by value of x

In both of these IR cases, the results are generally good. The predicted power law distributions are a decent fit for the actual observations: especially in the case of the alliances. Thus it seems that alliances and sanction sender coalitions follow power law distributions.

Discussion

This paper introduced the concept of power law behavior to various scenarios in political science. We found remarkable evidence of 'scale free invariance' in seemingly unrelated fields: Supreme Court Justice nominations, democide in comparative perspective and coalition formation at the international level. The 'real world' of empirical political science is generally messy, and it is exciting that I was able to find patterns behind seemingly irregular data. One imagines that, with more time, scholars could find other systems that obey these laws.

Undoubtedly, such work is well outside the mainstream of our discipline insofar as it does not represent the way we usually 'do business' in political science. There is no explicit modelling of human behavior here, either formally or informally. Rather, such work looks to a 'system' level of human interactions in their aggregate—an approach which has more in

common with theoretical physics (Cederman 2003).

But it is a mistake to assume that this inheritance from, and association with, another discipline is a shortcoming of the approach *per se*. Scholars may, for example, simply treat this paper as introducing a new set of distributions to our toolkit of statistical modelling to be used when appropriate with count data. Future work might treat α, the scaling parameter, as a quantity to be estimated from covariates in a way analogous to say, the p parameter in a binomial, the mean and variance of a normal and so on. Knowledge of power laws is not the same as *understanding* them, and this is where our discipline might now devote some attention.

References

Biggs, Michael. 2005. "Strikes as Forest Fires: Chicago and Paris in the Late Nineteenth Century." *The American Journal of Sociology* 110(6).

Cederman, Lars-Erik. 2003. "Modeling the Size of Wars: From Billiard Balls to Sandpiles." *American Political Science Review* 97:135–50.

Clauset, Aaron & Maxwell Young. 2005. "Scale Invariance in Global Terrorism.".
*http://www.citebase.org/cgi-bin/citations?id=oai:arXiv.org:physics/0502014

Gabaix, Xavier. 1999. "Zipf's Law for Cities: An Explanation." *Quarterly Journal of Economics* 114:739–67.

Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou & H. Eugene Stanley. 2003. "A Theory of Power Law Distributions in Financial Market Fluctuations." *Nature* 423:267–70.

Gliber, Douglas M. & Meredith Reid Sarkees. 2004. "Measuring Alliances: the Correlates of War Formal Interstate Alliance Dataset, 18162000." *Journal of Peace Research* 41(2):211–222.

Johnson, N, M Spagat, J Restrepo, J Bohorquez, N Suarez, E Restrepo & R Zarama. 2005. "From old wars to new wars and global terrorism.".
*http://www.citebase.org/cgi-bin/citations?id=oai:arXiv.org:physics/0506213

Mandelbrot, Benoit. 1997. *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk.* New York: Springer.

Martin, Lisa L. 1993. *Coercive Cooperation: Explaining Multilateral Economic Sanctions.* Princeton, NJ: Princeton University Press.

Richardson, Lewis F. 1948. "Variation of the Frequency of Fatal Quarrels with Magnitude." *American Statistical Association* 43:523–546.

Richardson, Lewis F. 1960. *Statistics of Deadly Quarrels.* Chicago: Quadrangle Books.